

**Univerzita Karlova v Praze**

Filozofická fakulta

Katedra sociologie

**DIPLOMOVÁ PRÁCE**

Petra Anýžová

**Teoretická a metodologická perspektiva mezinárodních komparativních výzkumů  
(Aplikace shlukové analýzy v srovnávacích analýzách European Social Survey)**

Theoretical and Methodological Perspective of Cross-national Comparative Surveys  
(Application of Cluster Analysis in Comparative Analyses of European Social Survey)

Praha 2010

Vedoucí práce: PhDr. Jiří Vinopal, Ph.D.

**Prohlášení:**

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

V Ústí nad Labem dne 20. dubna 2010

Petra Anýžová

**Poděkování:**

Velice děkuji svému vedoucímu diplomové práce PhDr. Jiřímu Vinopalovi, Ph.D. za ochotu, trpělivost a inspiraci, kterou mi při vedení práce poskytl, a za čas, který mi věnoval.

## **Abstrakt**

Diplomová práce se zabývá teoretickým a metodologickým zázemím komparativních mezinárodních výzkumů, jejichž význam v současné době nabývá na důležitosti, a to především z důvodů budování rozsáhlých kontinuálních mezinárodních programů realizující kvantitativní výzkumy se širokým okruhem sociálních témat, jakým je např. European Social Survey, do nichž se Česká republika po roce 1989 také zapojila. Prvním cílem bylo zmapovat přínosy, které komparativní studie přináší do propojování teorie a empirie na poli sociologie, a objasnit jejich schopnost přispívat k budování nových teorií a ověřování již vzniklých teorií. V centru pozornosti stál také popis specifických vlastností, jež komparativní studie oddělují od klasických národních studií. Druhým zájmem bylo popsat metodologické zázemí komparativních výzkumů s důrazem na naplnění konceptu ekvivalence výzkumu v mezinárodním kontextu a dosahování kvality mezinárodních šetření. Vedle obecného metodologického popisu byla pozornost věnována i specifickým přístupům k analýze sebraných mezinárodních dat, v jejichž prvních etapách empirické analýzy se jako nejvhodnější explorační statistický nástroj jeví využití shlukovacích metod.

### *Klíčová slova:*

komparativní výzkum, ekvivalence výzkumu, kvalita šetření, European Social Survey, shluková analýza

## **Abstract**

The thesis deals with the theoretical and methodological perspectives of comparative cross-national surveys. The importance of these perspectives currently increases, especially because of the creation of extensive continual international programs, as for instance European Social Survey, which realize quantitative researches with broad range of social topics. The Czech Republic began to engage in these international projects after 1989. The first goal of my thesis was to map the contributions which comparative studies bring into the combination of the theory and research in the area of sociology and illustrate the ability of comparative studies to contribute to the constructions of new theories and verifications of already existing theories. The thesis focuses also on the description of specific attributes which separate comparative studies from the standard monocultural studies. The second goal of my thesis was to describe methodological background of comparative survey researches with an emphasis on the concept of equivalence in cross-national research and on the achieving of quality of international surveys. Apart from the general methodological characterization an attention was paid to the specific approaches of the analysis of collected international data. Cluster analysis seems to be the most convenient exploratory statistical instrument in the first stages of empirical analysis.

### *Key words:*

comparative survey, equivalence in cross-national research, quality of survey, European Social Survey, cluster analysis

# Obsah

Seznam použitých zkratk	7
Úvod	9
1 Vymezení pojmu komparativní výzkum	15
2 Historie komparativních výzkumů	20
2.1 Příčiny rozvoje komparativních výzkumů a jejich přínos	22
3 Logika komparativních mezinárodních šetření	24
3.1 Teoretický přístup	24
3.1.1 Význam systémových faktorů v komparativních výzkumech	26
3.2 Metodologický přístup	28
3.3 Role kvantitativních metod v mezinárodních výzkumech	29
3.3.1 Výběr zemí do kvantitativního mezinárodního výzkumu	30
3.3.2 Sekundární analýzy v kvantitativních komparativních výzkumech	31
3.3.3 Metodologické problémy provázející kvantitativní mezinárodní výzkumy	32
4 Otázka ekvivalence v komparativních výzkumech	35
4.1 Reliabilita a validita měření	35
4.2 Multikulturní ekvivalence	37
4.2.1 Ekvivalentní měření ve výzkumu	38
4.2.2 Kategorizace ekvivalence	39
4.3 Techniky nastolování mezinárodní ekvivalence	40
5 Dosahování kvality v mezinárodních komparativních šetření	46
5.1 Statistický přístup k dosahování kvality šetření	48
5.1.1 Výběrový design	48
5.1.2 Chyby měření v komparativním výzkumu	51
5.2 Technicko-administrativní přístup ke kvalitě výzkumu	54
5.2.1 Metody sběru dat	54
5.2.2 Dokumentace komparativních výzkumů	56
5.3 Psychometrické přístupy v rámci dosahování kvality šetření	58
5.3.1 Způsob vytváření dotazníků a jejich překlad	58
5.3.2 Vývoj komparabilních otázek a odpovědí	60
5.4 Prostředky zmírňující problémy srovnávání (rady a doporučení)	62
6 Komparativní kvantitativní výzkum European Social Survey	65
6.1 Představení projektu The European Social Survey (ESS)	65
6.2 Metodologické standardy projektu ESS	67
7 Analýza mezinárodních dat v komparativních výzkumech	72
8 Principy a povaha shlukové analýzy	76
8.1 Klasifikace, typologie a taxonomie	76
8.2 Historie klasifikačních klastrovacích technik	78
8.3 Představení shlukové analýzy	79
8.4 Kategorizace klasických metod shlukové analýzy	82
9 Vlastní využití klastrovacích technik na datech European Social Survey	86
9.1 Hierarchická aglomerativní metoda	91
9.2 Metoda klastrování K-průměrů (K-means clustering)	110
9.2.1 Postup shlukování v metodě K-průměrů	110
9.2.2 Předpoklady a interpretace metody K-průměrů	112
9.3 Dvoukroková shluková analýza	118
9.3.1 Postup shlukování ve dvoukrokové shlukovací metodě	119
9.3.2 Interpretace výstupů dvoukrokové shlukové analýzy	122
9.4 Validita klastrování	128

10 Moderní klastrovací metody.....	131
Závěr.....	134
Seznam použité literatury.....	138
Seznam tabulek a grafů.....	146
Přílohy.....	148

## Seznam použitých zkratk

EB	Eurobarometer (Eurobarometr)
ESS	European Social Survey (Evropský sociální výzkum)
EVS	European Value Survey (Evropský výzkum hodnot)
ISSP	International Social Survey Programme (Mezinárodní sociální výzkumný program)
WVS	World Value Survey (Světový výzkum hodnot)

## Úvod

*„Research is always an adventure... But the adventure is even more hazardous when entering a field as new as cross-national research.“ [Christiansen 1951]*

Komparativní mezinárodní výzkumy představují skutečnou teoretickou a metodologickou výzvu v oblasti sociálních výzkumů. Tento široký termín zahrnuje kvantitativní i kvalitativní mezinárodní srovnávání více či méně rozsáhlých sociálních celků v různých kulturách, jejich geografických, politických, kulturních či ekonomických aspektů (např. nezaměstnanost, zdravotnictví atp.) a také komparaci mezi sociálními skupinami různých kategorií (sociální stratifikace, etnická stratifikace atp.). Komparace neodhaluje jen rozdílnosti mezi studovanými sociálními jevy, ale také ukazuje jejich unikátní společné aspekty, které by jinak nebylo možno nalézt a které činí z mezinárodních analýz nezbytnou část procesu rozvoje sociálněvědní teorie [Przeworski, Teune 1966: 555]. Záměrem komparace je nalezení sociálních vzorců nezávislých na čase a prostoru. Z tohoto důvodu tyto vzorce komparativní analýza odděluje z kontextu sociálního a kulturního prostředí a vnímá je jako univerzální zákonitosti. Takto může vědec také lépe zpětně porozumět specifickým charakteristikám sociálních jevů [Mills, Van de Bunt, De Bruijn 2006: 621]. Primárním cílem komparativních analýz je tedy oddělit od sebe obecný a kulturně specifický prvek ze zkoumaného předmětu.

Pokud chce sociologie naplňovat svou vlastní definici nomologické disciplíny, která se pokouší vybudovat obecné teorie sociálních fenoménů a procesů ve společnosti, musí zároveň nalézat tyto obecné zákonitosti v širokém časoprostorovém kontinuu celého sociálního světa. Teprve za těchto podmínek se mohou stát obecné výroky skutečně validní. Jejich empirické ověření by pak mělo správně využívat data z různých míst, kultur a historických momentů, tedy data z komparativních mezinárodních výzkumů, nejlépe kontinuálních. *„Sociologie tak musí být komparativní již ze své definice“.* [Nowak 1989: 34]

Pro výzkumníkovo osvojení zmíněné perspektivy sociologie slouží srovnávání jako nástroj k rozvíjení klasifikace sociálních fenoménů a pro zjišťování, zda sdílené fenomény mohou být vysvětleny stejnými příčinami. Pro mnoho sociologů srovnávání sociálních jevů ve dvou či více zemích poskytuje analytickou soustavu na přezkoumávání a vysvětlování sociálních a kulturních rozdílů a specifik. V poslední době je kladen velký důraz na tzv. kontextualizaci ve výzkumu, v tomto kontextu slouží mezinárodní srovnávání čím dál



více jako prostředek získávání lepšího porozumění odlišnostem společností, jejich struktury a institucím (zvyky, tradice, hodnotový systém, životní styl, jazyk, vzorce myšlení atp.) [Hantrais, Mangen 2006]. Účelem komparativních šetření tak může být vysvětlení podobností a rozdílností různých sociokulturních prostředí, jejich zobecňování nebo získávání hlubších znalostí o sociální realitě v různých národních kontextech.

V komparativních výzkumech se takto uplatňuje Dismanova [2005] myšlenka o tom, že sociolog téměř vždy studuje jiné sociální prostředí, než své vlastní, a tedy odlišnou kulturu, ve které je více či méně cizincem. A zatímco „domorodec“ každodenním situacím díky kulturním vzorcům rozumí, cizinec je musí racionálně analyzovat a vyložit – někdy i mylně, bez schopnosti rozlišit typické a individuální. Výzkumníci jsou v mezinárodních výzkumech vybízeni k akceptování různých kulturních pohledů, pokouší se tak rozumět sociálním procesům jiných kultur a vidět je s nadhledem. Během tohoto celého procesu je badatel zároveň nucen zpětně promýšlet svou vlastní národní kulturu z perspektivy vnějšího pozorovatele, což je zdrojem neocenitelných poznatků v oblasti národní vědy.

Z hlediska teoretických přínosů komparativních výzkumů je jejich význam zřejmý i v praktickém využití srovnávacích analýz, ve kterých se výzkumníci pokouší vytvořit obecné typologie vybraných sociálních systémů (zemí, států, společností, národů, kultur atp.). Nejčastěji se analýzy zabývají omezeným počtem sociálních systémů, u kterých lze sledovat stejné atributy (např. srovnání západoevropských demokracií) a nalézat nejružnější typologie na nižších úrovních (např. typologie totalitních režimů). Základem mnoha mezikulturních výzkumů bylo kupříkladu srovnávání moderních a tradičních společností [Przeworski, Teune 1970: 4]. Výzkumné otázky tohoto druhu komparativního výzkumu mají pak za předpokladu existence vývoje společností klasickou podobu: *Je například míra mobility v rozvinutých zemích menší nebo vyšší než v rozvojových zemích?*

Relevance tohoto tématu je patrná i v metodologické oblasti. Komparativní studie jsou více než jakékoli jiné studie metodologicky a strategicky hodnotné, nabízejí kvalitnější základ pro závažnější a systematictější úvahy o srovnatelnosti a ekvivalenci zásadních výzkumných témat v různých kulturních a sociálních prostředích a výzkumném designu takovýchto studií [Rokkan et al. 1969]. Jowell kategoricky ve svém článku „*How comparative is comparative research?*“ uvádí, že důležitost a užitečnost přesných a pečlivých mezinárodních měření sociálních výzkumů je neoddiskutovatelná [1998: 168].

Z metodologické perspektivy se vede zásadní debata o komparativních výzkumech jako specifické kategorii empirického sociálního výzkumu. Základní argumenty vycházejí

z faktu, že mezinárodní výzkumy doprovází řada metodologických problémů, které jsou jen těžko odstranitelné. Mezinárodní šetření se musí vypořádávat s různým významem pojmů, které odlišně interpretují různé sociální skupiny respondentů v různých kulturách. Totožný význam otázek a nabídek odpovědí v dotazníku není často zajistitelný ani stejnými měřicími nástroji použitými za stejných podmínek. Je až paradoxní, že některé ze zkoumaných odlišností mezi národy, jako je jejich rozdílné chování, životní poměry a postoje, by mohly ohrozit úspěch v dosažení standardizovaného přesného měření. Komparativní výzkum tak musí být předem navržený tak, aby byl aplikovatelný ve všech zkoumaných zemích a vznikaly z jeho sebraných dat kvalitní, podložené a věrohodné výstupy. Z tohoto důvodu se v oblasti komparativních výzkumů soustavně propracovává a vylepšuje výzkumná strategie, které se pokouší čelit všem zmíněným metodologickým problémům.

Někteří sociální vědci vidí tyto konkrétní výzkumné problémy jako nepřekonatelné pro kvalitní realizaci komparativních mezinárodních výzkumů, která je s ohledem na jejich teoretické ambice velice podstatná. Zvažují, že problémy jsou natolik závažné, že srovnávání mezinárodních dat se stává zbytečným, bezvýsledným nebo dokonce nepodloženým a nerozumným.

Na druhou stranu se staví ti vědci, kteří i přes výše zmíněné problémy vidí skutečné přínosy prováděných komparativních analýz. Věří, že měření mohou být promyšleně realizována tak, aby se zmírnil jejich dopad na srovnatelná data a uvolnil se prostor důležitým interpretacím výstupů srovnávacích analýz. Prosazují myšlenku, že pokud jsou výzkumníci z různých prostředí spojeni dohromady kvůli spolupráci na mezinárodních projektech, mohou být navázány cenné osobní kontakty, umožňující využít jednotlivých zkušeností a znalostí z různých kulturních tradic a srovnávat a hodnotit různorodost teoretických i výzkumných přístupů. Ale i výzkumníci podporující kladný názor o proveditelnosti kvalitních mezinárodních výzkumů důrazně podtrhávají možné chyby mezinárodního měření, vybízejí k větší opatrnosti v interpretaci výsledků analýz a navrhují řadu konkrétních opatření k nápravě a pomoci.

Rozvinutá mezinárodní spolupráce a dostatečné množství srovnatelných dat umožňuje v současné době přípravu nových komparativních šetření a prosazování nových témat. Nové zdroje dat je však potřeba znát, pracovat s nimi a dále je formovat. Zpravidla se jedná o otevřené a volně přístupné zdroje informací, které ale, mají-li mít smysl, musí být opřeny o solidně teoreticky a metodologicky založené koncepce. Podobně hodnota sociální teorie spočívá i v její empirické ověřitelnosti a rozvoj komparativních šetření k tomu otvírá jednu

z možných cest [Krejčí 2006: 151].

O důležitosti realizace mezinárodních šetření hovoří i skutečnost, že Mezinárodní sociologická asociace (ISA), která byla založena v roce 1949 a funguje dodnes pod záštitou UNESCO, věnovala už svou v pořadí druhou vydanou publikaci mezinárodnímu komparativnímu výzkumu. Není to náhoda, vedení ISA chtělo totiž už v padesátých letech 20. století výslovně upozornit na mimořádný význam, který má tato oblast výzkumů pro celý obor sociologie a který ovšem nebyl ještě léta náležitě doceněn [Krejčí 2006: 150].

O několik desítek let později Melvin L. Kohn, který se stal v roce 1987 prezidentem Americké sociologické společnosti (ASA), na každoroční kongres v Chicagu vybral jako hlavní pojednávané téma také „mezinárodní výzkumy v sociologii“. Kohn se ještě v 80. letech 20. století cítil jako misionář v této metodologické oblasti a jeho primárním cílem ve vedoucí pozici prezidenta ASA bylo vytvořit mezinárodní sociologii. Viděl tento směr jako výzvu pro americkou sociologii, kterou chtěl vymanit ze své provinciální perspektivy a vsadit Spojené státy do kontextu rozvoje ostatních zemí. V roce 1989 pak publikoval v díle „*Cross-national Research in Sociology*“ několik příspěvků, jež na zmíněné konferenci zazněly a které podle jeho soudu nejlépe vystihly možnosti a problémy komparativních výzkumů a způsob jejich využití. Až v této době se komparativní mezinárodní výzkumy dostávají do popředí zájmu mnoha sociálních vědců a ne jen ojedinělých skupin zainteresovaných výzkumníků.

V dnešní době se problematice komparativních mezinárodních měření systematicky věnují celá výzkumná oddělení důležitých akademických institucí, jedná se o frekventované téma předních odborných periodik z různých oblastí výzkumů, kterému se věnují též některé prestižní výukové programy a jsou na ně zaměřeny konferenční cykly [Krejčí 2008].

## **Koncepce práce**

Následující práce se snaží tyto přínosy komparativních výzkumů do efektivního propojování sociální teorie a empirie zachytit a vysvětlit jak z teoretické perspektivy, tak i z metodologického zázemí, čemuž odpovídá i struktura celého textu. První tři kapitoly se pokouší nastínit obecnou logiku kvantitativních komparativních výzkumů, jež jsou díky přítomnosti několika podstatných specifických vlastností vydělovány vůči národním šetřením. Sledují historický vývoj komparativních výzkumů, detailněji vymezují pojem komparativní výzkum, věnují se příčinám jejich masivního rozvoje a důležitým mezníkům změn jejich charakteristické povahy od srovnávání množství různorodých národních výzkumů z odlišných sociokulturních prostředích až k vytvoření důmyslných propracovaných a vysoce kvalitních

mezinárodních výzkumných programů.

Následující čtvrtá kapitola již zasahuje do metodologického zázemí komparativních výzkumů a snaží se vyzdvihnout podstatu otázky ekvivalence mezinárodních výzkumů a popsat a analyzovat techniky jejího nastolování v mezinárodním kontextu, navíc přidává i přístupy k odstranění problémů, které v rámci ekvivalence výzkumu pravidelně vznikají. Celý koncept ekvivalence propojuje pátá kapitola s nutností dosažení kvality v mezinárodních komparativních šetřeních, která ve výsledku určuje, nakolik budou výstupy analýzy mezinárodních dat interpretačně hodnotné. Všechny zásadní standardy kvality předvádí šestá kapitola na konkrétním příkladu realizace moderního mezinárodního výzkumného projektu European Social Survey, který v rámci vlastních kontinuálních sociálních šetření uplatňuje přísná metodologická pravidla.

Konečně poslední čtyři kapitoly se tématicky věnují analýze mezinárodních dat z komparativních výzkumů. Neocenitelným přínosem analýzy komparativních šetření je namísto pouhého kategorizování různých sociálních systémů a kultur systematická snaha nalézt vysvětlující proměnné, které mohou sloužit k odhalení mezinárodních odlišností a skrytých struktur v datech, a přispět tak k vytváření obecných sociálních teorií. Jak již bylo výše zmíněno, srovnávací analýzy slouží ke konstrukci obecné typologie sociálních fenoménů a sociálních systémů. Jejich hlavním záměrem je vysvětlování výskytu různých podobností a rozdílností sociokulturních prostředí, jejich zobecňování a získávání hlubších a akumulovaných znalostí o sociální realitě jako celku, ze kterých se systematicky vyvíjejí nové teorie a ověřují již existující teorie.

Těchto teoretických cílů komparativních analýz je možné dosáhnout, pokud bude do analýzy mezinárodních dat dosazena proměnná společnost (resp. kultura, země, národ, systém, stát atp.) jako jedna z nezávislých nebo kontrolních proměnných, která daný jev ovlivňuje; a pokud mezinárodní data projdou empirickým procesem statistické analýzy. Protože v případech mezinárodních šetření se jedná o zpracování značně velkého množství dat a proměnných, ideálním řešením je aplikace multidimenzionálních statistických technik. Pro teoretické potřeby srovnávacích analýz není většinou dostačující aplikace pouze jedné multidimenzionální techniky, jako nutné se spíše jeví jejich strategické propojení a promyšlené využití specifického analytického potenciálu každé z nich.

Neboť první kroky empirické analýzy mají spíše explorační ráz, zvláště pokud jde o tak velké množství jedinců a proměnných v analýze, jako nejlepší statistický nástroj se v této fázi jeví shluková analýza. Ta dokáže vytvořit kvalitní přehled v tak velkém množství

informací, nastínit nové teoretické typologie a efektivně pracovat s nezávislou proměnnou země jako klasifikační proměnnou pro tvorbu shluků podobných jedinců. Její teoretický přínos potřebám srovnávacích analýz spočívá ve schopnosti zachytit první signifikantní proměnné, které by mohly mít na systémové úrovni žádoucí vliv a vysvětlující potenciál pro různé skupiny objektů. Pokud se navržená typologie a zachycená struktura v datech zdá výzkumníkům signifikantní, dalším logickým krokem v procesu empirického zpracování mezinárodních dat je použití jiných multidimenzionálních technik (např. faktorová analýza, regresní analýza atp.), které již dokáží analýzu vztahů mezi proměnnými adekvátně prohloubit a systémové proměnné podrobně studovat.

Použití shlukové analýzy jako optimálního exploračního nástroje v komparativních analýzách se systematicky věnuje druhá část práce. Aplikace shlukové analýzy přitom v mezinárodních výzkumech skutečně není ojedinělá, což dokazují mnohé publikované odborné články, ve kterých se autoři pokouší využít všech možností, které shluková analýza pro prvotní přehled v datech nabízí. Takto byl její analytický přínos identifikován v oblasti rasové a etnické identity [DiStefano, Kamphaus 2006], rodinného chování, v politické oblasti [studie Wolfsona, Madjd-Sadjadi, Jamese 2004] i v oblasti ekonomického chování při nalézání industriálních shluků na základě koncentrované ekonomické aktivity [více studie Wolfa, Gertlera 2004].

Sedmá kapitola v tomto kontextu vysvětluje princip analýzy mezinárodních dat v komparativních výzkumech a přínosy shlukovacích technik do prvních fází empirického procesu statistické analýzy. Osmá kapitola se zabývá obecnými principy shlukové analýzy, jejím zázemím v klasifikačních metodách, stručným historickým vývojem a základní kategorizací klasických metod, které při analýze dat shluková analýza využívá. Devátá kapitola pak podrobně představí dvě klasické nepoužívanější shlukovací techniky v sociálních vědách (hierarchickou techniku a metodu K-průměrů) společně s třetí pokrokovější dvoukrokovou shlukovací technikou, která má za úkol nastínit rozšiřující se analytické možnosti moderních technik, které jsou stručně zachyceny v poslední desáté kapitole. Celý podrobný popis třech vybraných shlukovacích technik je doprovázen vlastními grafickými a numerickými výstupy z analýzy mezinárodních dat European Social Survey, které slouží jako názorná pomůcka při ilustrování analytických schopností klastrovacích metod v případě rozsáhlých dat mezinárodního výzkumu a pomáhají pochopit relativně technický popis metod.

# 1 Vymezení pojmu komparativní výzkum

Komparativní výzkum je velice komplexní a široký pojem. Komparativní šetření v sobě zahrnují výzkumníkovo rozhodnutí o objektu studia, geografickém rozsahu výzkumu, úrovni analýzy, ke které se odhodlá, i rozhodnutí o volbě komparace podle časového hlediska. Všechny tyto jednotlivé atributy pak ve výsledku činí jasné vymezení komparativního výzkumu značně obtížné. Snahu o příhodné vymezení ani neulehčuje množství užších nebo vágnějších definic komparativních a mezinárodních výzkumů, které se v odborné literatuře vyskytují, bez snahy o jakékoli komplexnější pojetí. Pokusme si tedy vymezit komparativní výzkum na základě následující klasifikace.

## A) Objekt výzkumu

Společnosti, které v komparativním výzkumu slouží jako objekty studia, mohou být ve výzkumu definovány různým způsobem jako země, státy, kultury či národy. Volba konkrétního objektu je velice různorodá a zdá se být i příčinou značné flexibility užívání pojmů komparativního i mezinárodního výzkumu bez bližšího specifikování. Tuto flexibilitu lze dokázat na nejednoznačném užívání cizojazyčného ekvivalentu pojmu komparativní výzkum. Zatímco totiž termín „*cross-national research*“ (mezinárodní výzkum) evokuje srovnávání zemí, států nebo společenských, kulturních a politických systémů, stejně tak často užívaný pojem „*cross-cultural research*“ (mezi-kulturní výzkum) hovoří spíše o komparaci kultur, etnik, národností a lokálních společností. V odborných publikacích lze vyzorovat také hojně zmiňování pojmů „*cross-societal research*“ (mezi-společenský, mezi-societální výzkum), který se svým charakterem blíží představě mezikulturního výzkumu, a „*cross-systemic research*“ (mezi-systémový výzkum), jež naopak zastupuje ideu mezinárodního výzkumu.

Různí výzkumníci upřednostňují různé objekty svého studia. Například Hantrais a Mangan [2006] poznamenávají, že definice zemí nebo společností mohou být problematické, protože to nejsou jednoduše identifikovatelné, trvalé a stabilní sociologické jednotky. Mohou mít odlišnou filozofickou či historickou podstatu a navíc ani nemusí odpovídat geografickému území národa. Z tohoto důvodu Hantrais a Mangan dále radí zaměřit mezinárodní výzkum na specifické organizační odvětví sociálního systému nebo zvažovat pouze subspolečenské jednotky namísto celé společnosti. Kohn [1989] zase na druhou stranu, z důvodu nejasného vymezení zemí, společností a kultur jako objektů komparativního studia, zvažuje jako jediný platný objekt národ. Ten má pro něj relativně jednoznačný význam a odkazuje k žádoucímu porovnávání systémů. Kohn navíc zdůrazňuje, že vědci musí jasně definovat, co je skutečným

objektem jejich studií, jaký druh empirického výzkumu v danou chvíli realizují, protože jinak redukuji kvalitu srovnatelnosti kulturních dat.

Fakt, že již zvolený konkrétní objekt nemusí sloužit jen jako pouhý objekt studia, ale může mít i další důležité role ve výzkumu, ukázal Kohn [1989: 21] ve své klasifikaci čtyř typů mezinárodních výzkumů. Národ, který může sloužit jako objekt studie, se může stát i kontextem dané studie, jednotkou analýzy i součástí většího sledovaného systému objektů.<sup>1</sup>

- ◆ *národ jako objekt studie* (studium konkrétní země či její sociální instituce)
- ◆ *národ jako kontext studie* (země funguje jako nástroj pro zkoumání sociálního kontextu závěrů výzkumu)
- ◆ *národ jako jednotka analýzy studie* (výzkum vztahů mezi charakteristikami dané země)
- ◆ *transnacionální charakter studie* (národ jako komponenta velkého internacionálního systému, studují se vzájemné vztahy mezi národy, resp. mezi objekty)

Objekt výzkumu může mít i užší vymezení, jedná se pak o srovnávání určitého fenoménu v rámci několika politických, kulturních či ekonomických systémů (volební systém, systém politických stran atd.). V případě užšího vymezení objektu studia je prvek analýzy lépe viditelný a jednoznačněji lze sledovat cíl, který si výzkumník vytknul. Tomuto pojetí odpovídá definice Hopkinse a Wallersteina [1967], kteří definovali mezinárodní výzkum jako analýzu statistických vztahů mezi charakteristikami jednotlivých zemí nebo agregátů na národní úrovni (např. příjem na jednoho obyvatele, etnické rozdělení, stupeň politické represe, průměrná životní očekávání atp.) [podrobněji Henkenrath 2002].

#### *B) Geografický rozsah výzkumu*

Sociologové začali používat termín „komparativní sociologie“ k označení výzkumu, jež překračuje limity jedné společnosti v jednom historickém okamžiku [Nowak 1989: 35]. Prostorová lokalizace je vedle časového umístění výzkumu podle sociálních vědců jedním z nejdůležitějších kritérií pro komparaci. Prostorovým umístěním se míní již pouhá příslušnost lidí k lidské společnosti nebo společnosti určitého typu (skupin, institucí atp.).

Komparativní výzkum je ale značně široký pojem, neboť se může týkat jak komparace v jedné konkrétní zemi (resp. společnosti), tak i komparace, která srovnává více zemí najednou. Podle Kohna [1989] je mnoho studií jednotlivých společností implicitně mezinárodních a komparativních, protože badatelé interpretují své závěry porovnáváním

---

<sup>1</sup> Existuje více takovýchto teoretických klasifikací např. Marsh 1967, Tilly 1984, Nowak 1989, Ragin 1989 atp. [podrobněji Kohn 1989].

znalostí o zemích, jež aktuálně teoreticky studují. I tyto srovnávací techniky uvnitř zemí (intra-cultural comparisons) se pak mohou jevit jako výhodný nástroj pro uchopení sociálních jevů. Srovnání se situací v jiných společnostech umožňuje odlišit významná pozorování od nevýznamných a existence různých referenčních skupin přibližuje použitou metodu experimentu a dodává výzkumu vyšší relevanci [Krejčí 2006: 150 – 151].

V případě, že je komparativní výzkum orientován na srovnávání více zemí najednou, je už považován za synonymní k mezinárodnímu výzkumu. Se širokou definicí komparativního výzkumu jako „*jakéhokoli výzkumu, který přesahuje národní hranice*“ [Przeworski, Teune 1966: 522] vzniká ale problém nedostatečného odlišení od zmíněné komparace v jednotlivých zemích, protože i v tomto případě badatelé přesahují národní hranice výzkumu tím, že porovnávají výsledky svého výzkumu s informacemi o jiných zemích. Kohn [1989: 20] z tohoto důvodu preferuje explicitní definici komparativních výzkumů: „*Studie, které systematicky používají srovnatelná data z dvou či více národů.*“ Toto pojetí postupem času převládlo a komparativní výzkum se již standardně definuje jako studium sociálních jevů ve dvou či více společnostech.

### C) Časové hledisko výzkumu

Kritérium časové lokalizace vymezuje komparativní výzkumy jako šetření srovnávající různá časová období. Z hlediska časové osy lze rozlišit tři základní varianty komparativního výzkumu.

- ◆ *Synchronní srovnávání* sledovaných fenoménů v různých společnostech nebo systémech v daném časovém úseku, který se v procesu komparace nemění a je stabilní. Tento typ srovnávání nezachycuje vývoj sociálního jevu, ale jeho stav v určitém historickém okamžiku. Jedná se o tzv. jednorázový výzkum.
- ◆ *Diachronní komparace* může zahrnovat srovnání jedné společnosti či jednoho stejného fenoménu v průběhu času (např. politický systém, studie hodnot). A nebo může zachytit vývoj více společností či více sociálních jevů v čase, které jsou mezi sebou porovnávány. Jedná se o opakovaný výzkum s jednou nebo více skupinami. Takto jsou realizovány mnohé mezinárodní komparativní výzkumy, které opakovaně po několika letech porovnávají desítky zemí. Tato tzv. kontinuální šetření považuje Krejčí [2006: 154] za nejnáročnější, avšak se širokou nabídkou možností mezinárodní analýzy. Kontinuální výzkumy zprostředkovávají zachycení trendů ve vývoji, hlubší analýzy postojů respondentů a sofistikovanější analýzy vztahů mezi vybranými



proměnnými.<sup>2</sup>

- ◆ Třetí variantou je *komparace analogických historických situací*. Jedná se o srovnávání podobné situace v různých částech světa (např. transformace nedemokratických systémů na demokratické). Analogických situací proto, že v nich probíhá stejný proces, ale v různých obdobích, v zemích s různými tradicemi, různým ekonomickým a politickým postavením atp. Tento typ výzkumu navozuje otázku, zda lze studii s takovýmto zaměřením provádět, resp. zda má hodnotu odpovídající vynaložené námaze. Na druhou stranu nelze tento typ výzkumu podceňovat. V případě, že je výzkumník schopen se vypořádat s výzkumnými i analytickými problémy tohoto typu srovnávání, mohou představovat tyto práce značný přínos ke klasifikaci politických a jiných systémů a ke sledování obecnějších trendů v jejich vývoji.

#### D) Úroveň zvolené analýzy

Przeworski a Teune [1970: 36] podotýkají, že ne všechny studie prováděné napříč zeměmi jsou komparativní. Komparativní výzkum je podle nich takové šetření, ve kterém je možná více než jedna úroveň analýzy a jednotky pozorování jsou identifikovatelné na každé z těchto vybraných úrovní. Všechny komparativní studie tak musí mít alespoň dvě odlišné úrovně analýzy – například individua a vyšší jednotku. Tato vyšší jednotka musí mít systémový charakter, být součástí sociálního, politického nebo ekonomického systému, proto se tato vyšší úroveň nazývá systémová a je to jakákoli úroveň vyšší než individuální. Pokud jsou v analýze zahrnuty jen individua a komunity, je tato systémová úroveň na úrovni lokálních komunit, pokud v analýze vystupují i národy, je systémová úroveň na úrovni národů. Výběr úrovní analýzy tak sahá od individuí, přes skupiny, komunity, instituce až k politickým systémům, národům a nadnárodním celkům.

Takto je kupříkladu studie lokálních vůdců vybraných z lokálních komunit v jedné zemi komparativní, protože výzkum je prováděn jak na individuální úrovni, tak i vyšší komunitní úrovni. Naproti tomu, pokud v analýze jedinců z různých zemí nebudou figurovat vyšší sociální celky (sociální skupiny, instituce atp.), výzkum komparativní nebude, i když jsou jedinci porovnáváni za různé země. Dokonce i když budou jednotky pozorovány ve výzkumu na několika úrovních, např. se výzkum zaměří na sledování individuí, komunit a zemí, a v analýze dat bude hrát roli pouze jedna úroveň, tak ani takovéto studie nemohou být podle Przeworského a Teuna [1970: 50] komparativní.

Primárním zájmem komparativních výzkumů je hledat systémové faktory, které

<sup>2</sup> Za kontinuální projekty pokládá Krejčí všechny, které se alespoň třikrát opakovaly a nebo jsou aktuálně založeny se záměrem dlouhodobého opakovaného výzkumu.

vznikají na zmíněné systémové úrovni. Tímto způsobem se komparativní studie snaží především sledovat vliv systému na individuální chování lidí. Aby mohl komparativní výzkum dospět k těmto cílům, volí si v analýze dat nezávislé proměnné (vzdělání, socializační proces atp.) ze systémové úrovně a závislé proměnné (politický postoj, identifikace s politickou stranou atp.) z nižší úrovně [Przeworski, Teune 1970: 49]. Pokud by se logika volby závislých a nezávislých proměnných v analýze obrátila a závislé proměnné by pocházely ze systémové úrovně a nezávislé z nižší úrovně, tak by studie sledovala vliv individuálního chování na chování systému. To však obvykle nestojí v popředí zájmu komparativních výzkumů.

## 2 Historie komparativních výzkumů

V empirickém období sociologických výzkumů od 20. do 50. let 20. století práce na komparativních výzkumech utichala, neboť rozvíjející se empirická orientace této vědní disciplíny usilovala o přímé, standardizované výroky o sociálních faktech. Tento trend vedl výzkumné otázky ke zjišťování vztahů mezi proměnnými uvnitř jedné společnosti, jedné kultury, jednoho sociálního systému, a závěry studií tak nepřekročily hranice národního státu. Vznikaly téměř výhradně monokulturní a monosystematické standardizované studie [Nowak 1989: 35].

Komparativní výzkumná orientace se opět vynořila po 2. světové válce. Sociologové začali používat termín „komparativní sociologie“ k označení výzkumu, jež překračuje hranice jedné společnosti. Výzkumy se staly mezinárodními, mezikulturními a srovnatelnými v různých časových (historických) momentech. Jejich výsledky měly významně přispívat nově se rozvíjející, precizně formulované, konzistentní, široce aplikovatelné a empiricky testovatelné sociální teorii [Nowak 1989: 36].

Až do konce 70. let většina mezinárodních výzkumů sestávala z behaviorálních a demografických studií (World Fertility Study, Multinational Time Use Studies, Demographic and Health Surveys atp.) a převážná většina těchto studií nebyla ještě v pravém smyslu komparativní (ani ve svém designu, ani v provedení a prvotním záměru). Byl to soubor národních dat, který byl re-kódován podle zvoleného principu kombinace informací z klíčových témat [Harkness, Mohler, Van de Vijver 2003: 4]. Nebo měly formu replikačních studií, které opakovaly výzkum z jedné země v nějaké jiné, pečlivě vybrané společnosti kvůli zjištění vlivu odlišných sociokulturních podmínek na zkoumaný fenomén [Straus 1968: 565].

Ve studiích tohoto typu se však kontinuálně objevovaly systematické i náhodné odchylky způsobené odlišnými sociálními organizacemi zkoumaných společností, odlišným druhem dotazování, různými výzkumnými nástroji a výběry, které znemožňovaly využít společnost jako nezávislou proměnnou v analýze a bránily systematickému srovnávání zemí [Madden 2004: 5-6; Hantrais, Mangen 2006].

Studie zemí, které byly součástí promyšleného designu rozsáhlých mezinárodních výzkumů, byly ojedinělé, a to především z důvodů finanční nákladnosti, organizační a technické složitosti a nejisté investice. Průlomovým byl v této oblasti výzkum Verby a Almonda *The Civic Culture* (1963), který se stal prvním pokusem o systematický sběr komparativních dat a systematickou mezinárodní analýzu na poli komparativní politiky a občanské kultury. Jako první sociální vědci navrhli Verba a Almond promyšlený výzkum

politického chování pro pět značně odlišných zemí a demonstrovali, že možnosti komparativních analýz nespočívají pouze v popisu odlišností a podobností mezi zeměmi (resp. politickými systémy), ale rozšiřují se až k empirickému testování obecných hypotéz. Verba tak dokázal podtrhnout hlavní rozdíly mezi monokulturním výzkumem a požadavky kladenými na mezinárodní výzkumy. I závažné analytické problémy, které se v jejich studii objevily, byly kritickou inspirací pro rozvíjející se mezinárodní výzkumy stejných ambicí [podrobněji Rokkan 1964].

Již na konci 60. let a počátkem 70. let se objevil nový druh strategicky promyšlených mezinárodních výzkumných projektů, podporovaných množstvím zákonodárných iniciativ v Americe i v Evropě. Jejich rozvoj souvisí s rozšířením informační technologie v 60. letech. Ještě ale v 70. letech probíhal v této oblasti především vývoj metod.<sup>3</sup>

Prvním příkladem je *Eurobarometer* (EB), který monitoruje sociální a politické postoje a hodnotovou orientaci dvakrát ročně v zemích Evropské unie (European Community) od roku 1973. Stejnou logiku sdílí výzkum *General Social Survey* (GSS), který se každoročně realizuje už od roku 1972 a sleduje strukturu a vývoj americké společnosti včetně mapování trendových dat postojů a sociálních změn v Americe.

V roce 1980 byl Světovou bankou iniciován vznik výzkumného projektu *Living Standards Measurement Study* (LSMS) z důvodů lepšího porozumění sociálním a ekonomickým vlivům v rozvojových zemích a také kvůli vylepšení kvality sbíraných dat z jednotlivých domácností. V roce 1981 započaly *World Value Survey* (WVS) a *European Value Study* (EVS) jako dlouhodobá komparativní šetření hodnot v sociální, kulturní a politické oblasti. Následoval v roce 1985 každoročně prováděný nejrozsáhlejší mezinárodní program *International Social Survey Programme* (ISSP), který se zaměřuje na postojové a behaviorální výzkumy. Na jeho počátku stály 4 státy, které se rozrostly až do počtu téměř 40 participujících zemí. V roce 1985 vznikla také skupina *Cross-National Research Group* s cílem poskytnout fórum pro diskuzi a výměnu názorů a zkušeností mezi výzkumníky z různých sociálněvědních disciplín zapojených do mezinárodních komparativních studií. Skupina dodnes pořádá semináře věnující se výzkumným metodám v mezinárodním kontextu, jejichž pořady koordinuje Linda Hantrais (Loughborough University UK), a také disponuje databází informací o výzkumnících zabývajících se mezinárodní komparativní prací.

V průběhu devadesátých let byl Eurobarometr rozšířen o celou další řadu dalších *Barometrů* (Latinobarometr v J. Americe, Afrobarometr, Barometr pro střední a východní

<sup>3</sup> Srovnávací výzkum v sociologii byl považován za ojedinělou oblast, která vygenerovala nejvíce metodologických pouček k nejmenší databázi a s nejméně výsledky [Kohn 1989: 78].

Evropu a Asiobarometr). Tato přelomová dekáda 20. století otevřela prostor pro realizaci mnoha dalších rozsáhlých, různě specifikovaných výzkumných šetření. Vznikají mezinárodní projekty, které mají podobu buď kontinuálních šetření nebo jednorázových výzkumů, mohou navíc vznikat v komerční sféře nebo v akademické oblasti. Příkladem jednorázového mezinárodního šetření je výzkum *European Quality of Life Survey* (EQLS) patřící pod program na podporu vytváření sociální politiky EU. Tento výzkum je zaměřený na otázky zaměstnanosti, příjmů a výdajů, vzdělání, bydlení, rodiny, zdraví, spokojenosti se životem atp. Šetření proběhlo v roce 2003 ve 28 zemích včetně ČR [Krejčí 2006: 164]. Příkladem kontinuálních šetření jsou projekty *OECD* zaměřené na vzdělávání, studium postojů k politickým stranám, monitoring veřejného mínění, příjmů domácností, pracovní mobility, zkušeností se zločinem atp. a nebo jeden z nejnovějších mezinárodních výzkumných programů *European Social Survey* (ESS), který provádí dotazníková šetření každé dva roky od svého vzniku v roce 2002. I Česká republika se zejména po roce 1989 zapojila do řady těchto mezinárodních projektů [stručný přehled poskytuje Krejčí 2006: 150].

Neméně významným zdrojem sociálních dat pro mezinárodní komparaci jsou oficiální statistiky. Nejdůležitější mezinárodní statistickou službou v rámci Evropy je Eurostat [Krejčí 2006: 166]. Eurostat zahájil v roce 2003 projekt *The European Union Statistics on Income and Living Conditions* (EU-SILC), který v roce 2009 získal účast již 31 zemí. Tento projekt poskytuje do Eurostatu každý rok kvalitní srovnatelné statistiky i longitudinální data ohledně příjmu, vzdělání, chudoby, sociální exkluze a životních podmínek. Dalším projektem přinášející podstatné informace do oficiálních statistik je rozsáhlý standardizovaný dotazníkový výzkum *Quality of working life survey*, který se realizoval od roku 1977 šestkrát. Má za úkol monitorovat pracovní podmínky zaměstnanců a jejich změny, pozici pracovníka na pracovním trhu atp.

## **2.1 Příčiny rozvoje komparativních výzkumů a jejich přínos**

K mohutnému rozvoji mezinárodních komparativních výzkumů podle Harkness et al. (eds.) [2003: 4] přispělo několik zásadních momentů. Mnohé politické bariéry sběru dat a jejich zveřejňování byly utlumeny v různých částech světa (nehledě na to, i dodnes je samozřejmě obtížné provést výzkum v některých oblastech světa, např. v Africe). Některá jiná teritoria své restriktce zcela zrušila. Celkově se zlepšil přístup k datům (resp. ochota jejich sdílení) a zesílily softwarové možnosti pro složitější analýzy. Výzkumníci, kteří se dříve

nemohli z finančních důvodů podílet na realizaci výzkumu, mají teď rozsáhlé možnosti výběru volně a snadno dostupných dat k širokému využití, vzrůstá tak počet publikací využívající sebraná komparativní data ke srovnávacím sekundárním analýzám.

K tomu přistupuje fakt, že objednávka společenskovedních informací založených na mezinárodních srovnáních roste a tím i možnosti realizace komparativních šetření [Krejčí 2006: 151]. Jejich potřebnost a aktuálnost vyvstává především z růstu globální integrace kulturních praktik států [Madden 2004: 4] a integrace nadnárodní politiky [Jowell 1998], včetně zvyšujícího se objemu mezinárodní komunikace [Krejčí 2006: 151], což činí analýzy cílených na mezinárodní srovnávací výzkum částečně nadřazené nad národně zaměřené analýzy [Madden 2004: 4]. Současné i vědy se stávají stále více mezinárodními a interdisciplinárními a široká spolupráce je dnes standardem [Hantrais, Mangen 2006; Krejčí 2006]. O zvyšování významu mezinárodních a mezikulturních výzkumů také svědčí přibývající počet specifických názvů pro multikulturní srovnávací vědní disciplíny, tj. komparativní sociologie, komparativní politické vědy, interkulturní komunikace, transkulturní psychiatrie, mezikulturní psychologie atp. [Harkness, Mohler, Van de Vijver 2003: 6].

Z množství dalších důvodů pro rozvoj srovnávacích kulturních statistik vystupuje do popředí snaha o dosažení lepšího, kvalitnějšího porozumění vlastní situaci (srovnáváním sama sebe s ostatními), poskytnutí globálního pohledu na danou problematiku, získání množství mezinárodních podnětů pro vytváření podrobnějších národních výzkumů. Výborně to shrnuje Kohn [1989: 77]: *„Pokud se zaměříme na téma výzkumné strategie, tak komparativní výzkumy nám z tohoto srovnání vyjdou s přemrštěnou cenou. Jsou časově i finančně nákladné, těžko realizovatelné, vyvolávají více interpretačních problémů než řeší. Jsou ale také teoreticky neocenitelné a podle mého názoru silně nevyužívané.“*

V kontextu české sociální vědy lze identifikovat další důvod k rozvoji mezinárodního výzkumu, kterým je jeho efektivita. S jistou finanční náročností se totiž váže několikanásobný přínos. Účast v mezinárodním projektu zprostředkovává pro české badatele přístup k zahraničním datům a dává možnost srovnávání. Přítomnost českých dat v mezinárodních souborech otevírá možnost zahrnout je do komparativních analýz i zahraničním vědcům. Podíl na organizaci šetření tak otevírá českým výzkumníkům přístup do mezinárodních struktur a zprostředkovává kontakt na mezinárodní vědu [Krejčí 2006: 171].

### 3 Logika komparativních mezinárodních šetření

Základní podstata a logika komparativních mezinárodních výzkumů je snadno zachytitelná porovnáním charakteru národních a mezinárodních výzkumů. To odkrývá překvapivé množství podobností i rozdílností mezi těmito druhy empirických šetření, na které se lze dívat dvěma různými pohledy. Za prvé, rozdíly mezi přístupy národních výzkumů a mezinárodních studií leží v teoretické oblasti. Přínos mezinárodních studií oproti národním spočívá především v jejich schopnosti konstrukce obecných teorií a výroků napříč společnostmi, jejich potenciál je spatřován ve shromažďování různorodých poznatků. Další podstatný rozdíl se ukrývá v soustavném hledání vlivu systémových faktorů na vysvětlovaný sociální jev, kterým se komparativní studie vyznačují. Teoretický přístup upozorňuje na to, že tyto přínosy komparativních výzkumů by neměly být opomíjeny.

Za druhé může být logika komparativních výzkumů zkoumána z metodologické perspektivy. Metodologický přístup k této otázce posuzuje specifické postavení mezinárodních výzkumů na poli empirických šetření a řeší řadu naléhavých otázek, které komparativní výzkumy doprovází. Rozdíl monokulturních a mezinárodních šetření je viděn především v kontextu specifických výzkumných problémů, se kterými se musí komparativní výzkumy vypořádávat (tzn. překlad dotazníků, školení tazatelů, zajištění kvalitních dat atp.)<sup>4</sup> Navíc se v mezinárodních výzkumech i jeho sekundárních analýzách převážně využívá možností kvantitativního výzkumu, který také řeší řadu metodologických otázek; všechny zmíněné body umisťují komparativní výzkumy na výjimečné místo v empirické oblasti sociálních šetření.

#### 3.1 Teoretický přístup

Jedním z předních zájmů sociálních vědců je budování obecných teorií lidského chování, bez ohledu na časoprostorové kontinuum. Jak podotýká Krejčí [2006: 150], přestože je definice sociologie nejasná, je v základu koncipována jako univerzalistická věda. Záměrem sociálních vědců je vysvětlování specifických událostí prostřednictvím výroků, které jsou obecně pravdivé od jednoho souboru životních podmínek ke druhým [Przeworski, Teune 1970: 18]. V podstatě se sociální vědci snaží z konkrétních událostí a jevů oddělit to, co je pro všechny zkoumané jevy společné.

Klíčový předpoklad, který doprovází rozvoj mezinárodních výzkumů již od počátku,

---

<sup>4</sup> viz. Johnson 1998; Jowell 1998; Harkness, Van de Vijver, Mohler (eds.) 2003; Madden 2004; Hantrais, Mangen 2006

je ten, že sociální mezinárodní výzkum by měl a může vést k vyřčení takovýchto *obecných výroků* o sociálním fenoménu [Przeworski, Teune 1970: 4]. Komparativní výzkum by měl mít schopnost vydělit z konkrétních událostí společné prvky, prvky, které budou platit pro všechny studované společnosti.

Obecné výroky, které jsou jedním z cílů komparativních výzkumů, by měly být optimální kombinací mezi výrokem přesným, a přesto dostatečně obecným; souhra kauzálního a zároveň stručného tvrzení [Przeworski, Teune 1970: 20]. To znamená, že musí vysvětlovat daný fenomén co nejpřesněji a nejkomplexněji v každém sociálním systému, na základě různých faktorů, jejichž počet je ale třeba omezit. Ovšem toto podrobné vysvětlení musí najít ideální rovnováhu s dostatečně obecným a rozsáhlým objasněním jevu. Při formulování obecné teorie musí být vysvětlující výroky otevřeny dalšímu rozšiřování o jiné případy. V praxi to znamená, že namísto toho, aby se v teoretických závěrech objevovaly názvy konkrétních společností (zemí), komparativní výzkum hledá vysvětlující proměnnou, která je schopna od sebe tyto společnosti odlišit.

Przeworski a Teune [1970: 75] uvádějí konkrétní příklady z výzkumu, ve kterém se zjistilo, že všichni mladí lidé manuálně pracující v průmyslovém závodě mají tendenci volit spíše levicové strany nehlédě na sociální systém, ve kterém žijí; tento poznatek platil pro Norsko, Francii, Chile atp. Pokud se do analýzy přidala ke zkoumání i proměnná pohlaví, zjistilo se, že pro tento vztah platí, že ve Francii a Chile tendence volit levici závisí na pohlaví a že větší tendenci volit levicové strany mají muži než ženy. V Norsku už však toto rozdělení podle pohlaví neplatilo, tam se ukázala tendence volit levici nezávislá na pohlaví. Pokud se do analýzy přidala proměnná role náboženství v zemi, která už fungovala na systémové úrovni, vědci zjistili, že rozdíly mezi pohlavím na této strukturální proměnné silně závisí. Pokud je role náboženství v zemi silná (Francie), vznikají rozdíly ve volebním chování podle pohlaví, pokud náboženství v zemi nehraje tak důležitou roli (Norsko), je volební chování na pohlaví nezávislé. Vidíme, že mezi volebním chováním mužů a žen vznikají systémové rozdíly, které jsou značně podstatné.

Touto systémovou proměnnou může být chování Francouzů i Norů vysvětleno stejnými teoretickými pojmy, a přesně tento cíl chtějí komparativní výzkumy naplnit. V této souvislosti mohou být všechny zahrnuté země v analýze nahrazeny pojmem vysvětlující systémové proměnné: „*sociální systémy, ve kterých náboženské organizace buď hrají nebo nehrají důležitou roli*“. Namísto specifikování názvu sociálních systémů (společnosti, země, národu) je důležité do teoretického objasnění zařadit vysvětlující proměnnou, která funguje



na úrovni odpovídající sociálnímu prostředí a která konkrétní název systému nahradí. Názvy sociálních systémů v obecné teorii pak vystupují jako souhrn teoretických proměnných, které od sebe byly schopny tyto sociální systémy odlišit [Przeworski, Teune 1970: 75]. Výroky o vztazích proměnných, které jsou nezávislé na konkrétním pojmenování systémů zemí a jsou tak validní ve všech sociálních systémech, jsou obecné výroky, kterými komparativní výzkumy významně přispívají k budování sociálněvědních teorií.

Vidíme, že pokud je komparativní studie schopna vytvářet obecné výroky, může pak komparativní výzkum sloužit jako *nástroj k systematické akumulaci znalostí o sociální realitě, k potvrzování (testování) a modifikaci teorií*, které vznikají v různých sociálních prostředích [Przeworski, Teune 1970: 86]. Zároveň jsou srovnávací studie výzvou k interpretaci historických událostí, protože i když jsou historické události unikátní a vázané na sociální kontext dané doby, ještě to neznamená, že jejich vysvětlení nemusí být založeno na obecných teoriích.

Se změnou povahy sociální vědy od 70. let přichází viditelný důraz na propojování teorie s výzkumem a budování teorie samotné [Przeworski, Teune 1970]. Pokud je komparativní výzkum natolik provázán s vytvářením teoretických obecných výroků, je jeho pokročilý vývoj v této době zcela logický. Navíc dokáže data zachytit uvnitř sociálního kontextu, což je dnes také žádaná vlastnost výzkumu. Od 70. let vzrůstá i možnost dosažení daleko většího teoretického poznání spoluprací mnoha schopných sociálních vědců z celého světa, než jakého by byl schopen každý zvlášť v národní studii. Přirozeně se tak otevírá prostor pro početnou realizaci komparativních studií, které danému účelu nejlépe poslouží.

### **3.1.1 Význam systémových faktorů v komparativních výzkumech**

V komparativních výzkumech existují na rozdíl od národních šetření vždy minimálně dvě úrovně, které je nutné analyzovat. Jedná se o individuální úroveň, která má podobu zkoumaného sociálního fenoménu, například odlišnosti mezi lokálními vůdci, socializace v rodině, obavy obyvatelů země z příchodu imigrantů atp. Druhou úrovní je vyšší systémová úroveň, která odkazuje k faktorům, které mohou zkoumaný fenomén ovlivňovat, například velikost města, která může mít vliv na zmíněnou socializaci v rodině, nebo ekonomická situace země, která může mít vliv na obavy z příchodu imigrantů.

Výchozím předpokladem mezinárodního komparativního výzkumu je potenciální vliv tohoto systémového faktoru na vysvětlovaný fenomén [Przeworski, Teune 1970: 40; Nowak 1989]. Pokud sociální systém nebo společnost zařazená do analýzy nepomáhá charakterizovat

a pochopit individuální jednání a postoje lidí, systémové faktory nejsou ve výsledku důležité. Výzkum jasně nerozlišuje charakter jednotlivých sociálních prostředích. Jestliže se neliší orientace lokálních vůdců na sociální změnu (např. revoluci) v Polsku, Indii, Jugoslávii a Americe (tzn. ve velice odlišných zemích), můžeme očekávat, že se nebude signifikantně lišit ani v jiných zemích a že systémové faktory nejsou důležité pro vysvětlení tohoto konkrétního postoje [Przeworski, Teune 1970: 42].

Stejně tak lze zvažovat systémové ovlivnění teoretických vztahů proměnných v různých sociálních prostředích. Země se liší kupříkladu svou úrovní vzdělání, třídní strukturou a rodinnou socializací, ale neliší se jako systémy pokud jsou jejich struktury vztahů mezi proměnnými stejné. V centru pozornosti komparativních studií je vždy systematické pozorování těchto nezávislých proměnných na systémové úrovni, resp. vlivu systému na individuální chování uvnitř systému [Przeworski, Teune 1970: 49, 57]. Podle Przeworského a Teuna [1970] lze nalézt tři druhy systémových faktorů, které mohou mít potenciálně vliv na individuální chování lidí:

- ◆ *Difúze (Diffusion Patterns)*: podobnost individuálních chování a postojů ve skupině stejných zemí může být způsobena šířením kulturních vzorců [Naroll 1965]. Z tohoto poznatku následně plyne podle Przeworského [1970: 53] zásadní otázka: Dějí se sociální změny z důvodu proměn odehrávajících se uvnitř systému nebo protože byla země vystavena hodnotám a chováním jiných společností? Je to sociální změna nebo sociální absorbce?
- ◆ *Prostředí (Settings)*: vlivu prostředí jsou vystavena všechna individua uvnitř systému, a tento vliv nelze zkoumat na individuální úrovni. Przeworski [1970: 54] vyděluje prostředí historické, institucionální, externí (vnější vlivy systému), behaviorální (chování jedinců) a fyzikální prostředí.
- ◆ *Kontext (Contexts)*: kontextové faktory mohou mít systémový charakter (třídní struktura, nerovnost příjmů, komunikační síť) a individuální charakter (osobnostní charakteristiky lidí)

Toto ovlivnění sociálním prostředím je závislé na povaze sledovaného fenoménu, psychologický jev bude pravděpodobně mnohem méně závislý na sociálním systému než politický jev [Przeworski, Teune 1970: 42]. Některé psychologické teorie dokonce tvrdí, že rozdíly osobností uvnitř kultury jsou daleko větší než rozdíly osobností mezi kulturami [Singer 1961]. I sociologické studie 50. a 60. let, věnující se hodnocení prestiže zaměstnání v různých společnostech, nakonec z výzkumů vyvodily obecné závěry o tom, že lidé jsou si

názorově daleko vzdálenější uvnitř jedné země, než mezi různými zeměmi (viz. Inkeles, Rossi 1956). Sociálněvědní teorie možná jen poněkud zveličují odlišnosti mezi společnostmi a roli systémových faktorů a úloha komparativních výzkumů pak leží ve vyvrácení těchto mýtů, neboť jsou schopné tyto obecné systémové faktory zachytit [Przeworski, Teune 1970: 42].

### 3.2 Metodologický přístup

Z metodologické perspektivy se vede zásadní debata o komparativních výzkumech jako specifické kategorii empirického výzkumu. Na jedné straně stojí stanovisko prosazující totožnost monokulturních a mezinárodních sociálněvědních výzkumů. Z teoretické pozice argumentují tím, že v podstatě všechny výzkumy jsou komparativní a vzájemně srovnávají sociální jevy. Z metodologické oblasti vysvětlují, že národní a mezinárodní výzkumy se sice odlišují cenou a obtížností realizace, ale pravidla statistické indukce se v analýzách používají v obou případech stejně. Zkoumání reliability položek je také stejně relevantní pro národní i mezinárodní výzkumy. Navíc si mohou národní výzkumy záměrně přizpůsobit design výzkumu libovolně podle konkrétní výzkumné situace, což z nich činí flexibilnější analytické nástroje, ze kterých je vhodné čerpat inspiraci [Harkness, Mohler, Van de Vijver 2003: 8]. Tato strana upřednostňuje v rámci komparativních šetření přijetí totožných principů metodologie, jakou využívají národní výzkumy.

Na druhé straně stojí opoziční protiargument, který se snaží dokázat existenci jistých metodologických otázek, jež vyvíjejí tlak na komparativní výzkum a odlišují ho tak od monokulturních studií. Zásadní postřehy se týkají faktu, že komparativní výzkum by měl být navržený a předem určený pro implementaci do různých kulturních skupin. V mezinárodních šetřeních totiž nemáme žádné důkazy k předpokladu identity významů pojmů mezi různými sociálními, lingvistickými a kulturními skupinami. Navíc v odlišných kulturních skupinách není totožná interpretace otázek a kategorií na škále odpovědí zajištěna ani při použití stejných měřicích nástrojů ve stejných podmínkách. Vědci proto musí v komparativních studiích už předem pečlivě zvážit zahrnutí specifických výzkumných položek do dotazníku, použití konkrétních výzkumných nástrojů, metod sběru dat, přístup k překladům dotazníků. Měli by i obezřetně přistupovat k samotnému výběru zemí do výzkumného programu, aby náročné vypořádávání se se všemi problémy nezastínilo skutečný přínos a efektivitu mezinárodního šetření.

Z tohoto důvodu se v oblasti komparativních výzkumů soustavně pracováá a

vylepšuje výzkumná strategie, které se pokouší čelit faktu, že teoretické koncepty nemusí být identické nebo komparabilní a že zvolené instrumenty měření (resp. výběrové metody, techniky měření, sběr dat, procesy analýzy) mohou být různě adekvátní v různých kontextech odlišných kultur [Harkness, Mohler, Van de Vijver 2003: 8]. Design, nástroje měření i strategie výzkumu se vytříbují podle principu universálnosti s ohledem na cíle výzkumu, základní principy kvality výzkumu, aplikaci designu i následnou realizaci výzkumu. Nejběžnější přístup, který se pro realizaci komparativního výzkumu volí, je vybrat si design výzkumu a replikovat ho tím nejlepším možným způsobem v každé participující společnosti (tzv. best practise principles).

Předpoklad o jednoduché přenositelnosti technických standardů výzkumu z monokulturní oblasti do mezinárodní oblasti výzkumu bez zvažování všech důsledků se v tomto kontextu jeví jako kontraproduktivní. Krejčí [2006: 171] ale trefně shrnuje, že pohled na národní a mezinárodní výzkum jako na dvě odlišné oblasti a konkurenty je lichý. Logika komparativních výzkumů se nijak neliší od jiných typů sociálněvědních výzkumů, pokud studujeme například vztah mezi pohlavím a politickou orientací, ať už mezinárodně nebo pouze v jedné zemi, logika výzkumu zůstává stejná [Przeworski, Teune 1966: 522; Hantrais, Mangen 2006]. Východiskem pro formulaci priorit a koncepce výzkumu by vždy měl být konkrétní výzkumný problém, potřeba odpovědi na otázky a nikoliv metoda. V tomto smyslu by účast v mezinárodních šetřeních neměla vést k vytváření samostatné a specifické oblasti výzkumu [Krejčí 2006: 171].

S tím souvisí čím dál více verbalizovaná potřeba změny vědních disciplín, volajících po spolupráci a dvousměrném přenosu výzkumných znalostí a standardů z monokulturní do mezinárodní komparativní oblasti sociálního výzkumu a naopak (viz. sociálně kognitivní přístupy) [Harkness, Mohler, Van de Vijver 2003: 6].

### **3.3 Role kvantitativních metod v mezinárodních výzkumech**

V současné době komparativní metody pokrývají široké pole kvantitativních (variable-oriented approach) i kvalitativních (case-oriented approach) výzkumných metod. Vede jednotný názor v tom, že zde neexistuje žádná výsostná privilegovaná metoda. Avšak stále jedním z nejdůležitějších a nejběžnějších zdrojů faktických informací pro mezinárodní srovnání je *systematický kvantitativní sociální výzkum* na úrovni jedinců a na národní úrovni (tzv. makrosociologický výzkum). Bollen, Entwistle a Alderson [1993: 322] v roce 1993

napočítali v nejpřednějších sociologických časopisech American Journal of Sociology, American Sociological Review a Social Forces více než 80 článků o mezinárodních výzkumech, ve kterých mezi roky 1985 a 1990 aplikovali vědci kvantitativní metody, a našli více jak 200 odborných sociologických publikací makrokvantitativního zaměření zařazených v tématické sekci Contemporary Sociology.

### **3.3.1 Výběr zemí do kvantitativního mezinárodního výzkumu**

Když na konci 80. let Ragin [1989] bilancoval stav komparativního výzkumu, upozornil, že existuje relativně hodně studií srovnávajících dvě nebo několik společností do větší hloubky a množství povrchnějších srovnání velkého počtu zemí (třiceti a více), ale jen málo prací mezi tím. Tato skutečnost byla podle Krejčího [2006: 151] důsledkem charakteru dat, které byly k dispozici. Na jedné straně bylo možné provádět šetření s limitovaným rozsahem a zakládat je na jiných než kvantitativních metodách, na druhé straně byla dostupná data ze statistik Světové banky a dalších organizací (tzv. oficiální statistiky) zachycujících jen úzký průřez reality, ale v mnoha zemích.

Současná sociální věda (včetně komparativního výzkumu) se snaží překlenout rozdíl mezi extenzivním přístupem (např. rozsáhlá kvantitativní šetření) a intenzivním přístupem (např. případové studie, narativní přístupy) [Ragin 1989: 61]. Na místo prosazování jen jednoho přístupu se zabývá spíše teorií středního dosahu, která funguje jako prostředník mezi těmito odlišnými pojetími. Teorie středního dosahu nezamýšlí popsat náš sociální systém jako celek, ale spíše jako jeho podskupinu, rozvíjet sociální teorie, jež by se daly použít na vysvětlení omezeného množství jevů, a takto se pokusit dát nový směr empirickému zkoumání. Jako názorný příklad tohoto přístupu může sloužit komparativní šetření, které pátrá po rozdílnostech ne třiceti zemí, ale dvou či několika sociálních systémů, a který zjištěné rozdíly mezi zeměmi srovnává ve vztahu k různým proměnným tak, aby stanovil, zda a jakým způsobem spolu souvisí [Przeworski, Teune 1970: 31].

V tomto kontextu se objevuje otázka vhodného výběru zemí jak do primárního mezinárodního výzkumu, tak do sekundárních analýz. Nejčastěji se při výběru počtu zemí argumentuje faktem, že neexistuje přímá souvislost mezi množstvím zemí zahrnutých v analýze a množstvím relevantních získaných informací [Kohn 1989: 95]. Vždy jde o kompromis mezi počtem studovaných zemí a kvantem získané informace a toto rozhodnutí leží na výzkumníkovi. Beze sporu ve sběru primárních dat i v sekundárních analýzách může být důležitá výhoda v konzistentních a rozsáhlých výstupech napříč mnoha národy, kulturami

i politickými systémy. Výběr zemí by ale měl být vždy determinován otázkou, zda se srovnávání těchto konkrétních národů vyplatí i za cenu investovaného času a financí, které mezinárodní výzkum zcela jistě vyžaduje [Kohn 1989: 96]. Země by měly být vyselektovány na základě podobných historických a politických procesů, aby jako odlehlá pozorování nezkreslovaly výsledky [Mills, Van de Bunt, De Bruijn 2006: 622]. Navíc směs zemí vybraných v komparativních studiích ovlivňuje kvalitu a komparabilitu dat stejně jako povahu spolupráce mezi výzkumníky.

Kohn [1989: 95] preferuje strategii promyšleného výběru malého množství národů pro systematickou intenzivní studii, kde mají zvolené národy nejsilnější vztah k testovanému teoretickému jevu. Navrhuje tak začít studií jedné země a poté rozsah výzkumu rozšířit i do jiných oblastí. Ať už je ale vybrán jakýkoli postup výběru, vždy bude mít své výhody a nevýhody, vždy bude proti sobě stát celistvost výsledků z výběru jen několika vhodných zemí a možnost generalizace výroků z velkých šetření [Ragin 1989].

### **3.3.2 Sekundární analýzy v kvantitativních komparativních výzkumech**

Sekundární analýzy kvantitativních dat jsou v komparativních výzkumech relativně rozšířené, nepochybně částečně z důvodu ceny a časové náročnosti získávání primárních dat. Van Deth [2003: 291] v souvislosti s prosazováním plošného rozšiřování sekundárních analýz, jakožto kvalitního analytického nástroje, poukazuje na přeceňování potřeby nově sebraných dat v současné sociální vědě: „*Představa o tom, že sebrat svá vlastní data, je ideální situací pro výzkumníky, je založena na zřejmém neporozumění role empirického testování a explorační ve výzkumu.*“

Hyman [1972] vymezuje sekundární analýzu jako produkování znalostí založených na jiném tématu než na tom, na který se soustředil původní výzkum. Sekundární analýza může nabývat podoby verifikačního výzkumu, který zodpovídá nové otázky starými daty, nebo replikačního výzkumu, jež adekvátně rozšiřuje dostupná data doplňujícími závěry a zvyšuje věrohodnost prvotního výzkumu tím, že opakovaně testuje jeho reliabilitu.

Pro komparativní výzkum je práce s již sebranými daty zcela zásadní. Obzvláště v případě, pokud je výzkumníkovým cílem dosáhnout informací o rozvoji sociálního fenoménu, může tak na základě sekundárních analýz podávat komplexní studie o sociálních změnách a dynamických procesech ve společnosti [Van Deth 2003: 293].

Sekundární analýzy mezinárodního charakteru nacházejí nové interpretace dat, agregují datové analýzy a srovnávají velké množství dat (převážně z vládních zdrojů),

obzvláště náročné výzkumy v podobě časových řad (tzv. longitudinální analýzy) se prakticky neobejdou bez již sebraných dat. Z hlediska získávání těchto dat do sekundárních analýz mají být základním nástrojem pro akademický komparativní výzkum buď národní datové archivy, jejichž překážkou je ale nízká úroveň propojení datových knihoven vlastních každému datovému archivu. Některé důležité komparativní databáze jsou dostupné ve většině národních archivů (ISSP, EVS), jiné jsou ale uloženy pouze na jednom nebo několika místech [Krejčí 2006: 152]. Jiným zdrojem mohou být komparativní databáze budované na internetu jako vlastní výstupy řady mezinárodních výzkumných projektů (ESS). Tyto databáze kvantitativních šetření jsou veřejně dostupné pro sekundární zpracování a tvoří tak základní infrastrukturu pro sekundární výzkum [Krejčí 2006: 149].

Principiální problémy sekundárních analýz mezinárodních dat se pojí s nejrůznějším druhem překážek, které znemožňují opětovnou analýzu již sebraných dat. Ty lze najít ve třech metodologických oblastech a to odlišné operacionalizaci položek, nevhodném měření klíčových jevů a neadekvátní dokumentaci [Van Deth 2003: 296]. Precizní dokumentace sebraných (archivovaných) dat publikovaná společně s datovým souborem má zahrnovat popis výběrové procedury, design měřících nástrojů, detaily terénního sběru dat a popis finální podoby datového souboru. V této podobě pak umožňuje kvalitní a validní analýzu.

Odlišnou operacionalizaci a nevhodné měření klíčových teoretických pojmů řeší sekundární analýzy v procesu tzv. harmonizace měřících nástrojů. Harmonizace instrumentů z různých výzkumů vybraných pro konkrétní sekundární analýzu znamená v podstatě rekódovat data do optimálního kódovacího schématu [Van Deth 2003: 297]. Například se překóduje celá struktura hodnocení pracovních pozic ve všech zemích, protože v primárním výzkumu byla odlišná kódovací schémata pro tuto proměnnou výsledkem kompromisu výzkumníků, kteří tak řešili kulturní a historické rozdíly ve významech pojmů z důvodu budování ekvivalence [Van Deth 2003: 302].

### **3.3.3 Metodologické problémy provázející kvantitativní mezinárodní výzkumy**

Použití metod přímého statistického porovnávání kvantifikovaných údajů s sebou nese vedle počtu skutečných přínosů, kterým je především cenová i časová efektivita pro sekundární analýzy čerpajících z bohatého informačního zdroje, i řadu metodologických problémů.

V první řadě figuruje v makrosociologickém výzkumu otázka sporné komparability politicky, ekonomicky a informačně provázaných národních společností. Existence různých

regionálních dynamik a přenosu sociálních fenoménů v provázaných společnostech může nakonec podkopat nezávislost pozorování v mezinárodních výzkumech. Tato kulturní propletenost zemí představuje specifický problém ve statistické analýze z důvodů vytváření vnějších vlivů, které mohou působit pouze na jednu část národního vzorku [Herkenrath 2002: 522]. Tento jev byl v roce 1889 na setkání Royal Anthropological Society nazván „*Galton problem*“ podle Sira Francise Galtona. Proti tomu stojí argument Goldthorpa [1997], který prosazuje, že tyto vlivy nemají vždy významnou roli a že vzájemné propojení mezi zeměmi je problematické pouze, pokud zůstává skryto a nepozorováno. Podle něj může být tento jev zahrnut do designu kvantitativního výzkumu bez vážnějších problémů s metodologickou chybou.

Za druhé prosazování striktně kvantitativního přístupu může přivést sociální vědy k jednostrannému empiricismu. Se vzrůstající zásobou mezinárodních kvantitativních dat se totiž čím dál více uplatňují v analýzách několika úrovněvých statistické modely (*multilevel models*), které studují proměnné na různé úrovni agregace od individuální úrovně až po národní úroveň. Například školní úspěch žáků může být studován jako funkce jejich osobní inteligence, pedagogických schopností učitele, kvality školy, peněz vynaložených státem do vzdělávání atp.

Některá agregovaná data mohou být shromážděna pouze na nadindividuální úrovni. Pokud analýza takovýchto agregátních dat slouží jako nástroj pro studium chování jedinců, mohou se na základě toho produkovat nesprávné závěry a dojít k chybě zvané *ecological fallacy* [Przeworski, Teune 1970: 72; Herkenrath 2002: 520; Kostecký, Čermák 2003: 532; Van de Vijver 2003b: 223]. Statistický vztah významný na agregátní úrovni nemusí být totiž významný na úrovni jednotlivce. Robinson (1950) byl schopný nalézt spousty reálných příkladů, kde tato chybná interpretace vzniká. Například situace dobrých žáků ve špatné škole, na něž jsou vztahovány školní charakteristiky, a jsou tudíž pokládáni za špatné žáky [podrobněji Van de Vijver 2003b]. Největší komplikaci při analýze agregátních dat představuje skutečnost, že jednotlivé charakteristiky, které slouží v agregátních analýzách jako vysvětlující proměnné, jsou často silně vzájemně zkorelovány a je u nich obtížné rozhodnout, co je příčinou a co je následkem [Kostecký, Čermák 2003: 532]. Opakem zmíněného jevu je *individual-differences fallacy* představující chybné názory z přenesení fenoménu na individuální úrovni do agregovaných jednotek, například předpoklad, že země s velkým počtem chudých lidí má nízké HDP [Van de Vijver 2003b: 224].

Výslovnou kritiku principů makro-kvantitativního komparativního výzkumu a



agregovaných analýz shrnula Nancy Cartwright (2002) do provokativního pojmu „*crazy methodology*“, kterým nazvala makrosociální výzkum. Obvinila tento typ empirického výzkumu z napodobování přístupu přírodních věd a posoudila předpoklady homogenity, nezávislosti, případné linearitě závislostí v agregovaných analýzách jako zcela nevhodné pro potřeby sociálních věd [podrobněji Kittel 2006]. Kittel kritiku Cartwright nepatrně zmírňuje a hovoří o omezeném porozumění sociálním makro-fenoménům a o značných limitech makro-kvantitativního výzkumu, kterých si musí být výzkumníci vědomi [2006: 666]. Sociální jevy jsou totiž na makroúrovni vzájemně propojené a příliš závislé na specifických sociálních prostředích, ve kterých se vyskytují, než aby jejich sledování vyprodukovalo obecné výroky. Na takto vysokém statistickém stupni výzkumy provází nedostatečná robustnost a častá přítomnost příliš obtížných analýz náchylných k nepřesnostem a matematickým konstrukcím (tzv. *tricky analysis*).

Řada specifických problémů existuje u mezinárodních komparativních výzkumů i v oblasti interpretace a explanace výsledků. Jejich multidisciplinární přístup je výhodný svou flexibilitou, kterou nabízí, ale přesto stojí komparativní studie často tváří v tvář otázce, zda jejich výzkumníci nepostrádají ve svých výstupech souvislý jednotící celek. Zejména tedy obtížnost mezinárodních výzkumů vede až k určité specializaci uvnitř sociologie.

Jak již bylo v předchozích kapitolách zmíněno, na rozdíl od národních výzkumů jsou kvantitativní komparativní výzkumy vystaveny především tlaku metodologických otázek srovnatelnosti a adekvátnosti použití stejných měřících položek a škál odpovědí v dotaznících a problematice aplikace odlišných výzkumných strategií v různých zkoumaných zemích. Z tohoto důvodu se v oblasti komparativních mezinárodních výzkumů otevírá největší metodologický problém ekvivalence výzkumu, kterému je věnována následující samostatná kapitola.

## 4 Otázka ekvivalence v komparativních výzkumech

### 4.1 Reliabilita a validita měření

Standardní definice ekvivalence zahrnuje do svého obsahu termíny jako „*rovnocennost pojmů, rovnocennost významu pojmů, vzájemná korespondence pojmů, někdy i faktická identita pojmů, identický vliv položek a stejná funkce položek*“ [Johnson 1998]. V mezikulturních a mezinárodních šetřeních se ekvivalence výzkumných otázek rovná svou důležitostí jejich reliabilitě a validitě, proto je jí věnována tak velká pozornost. V následující kapitole bude vysvětlena celková provázanost těchto důležitých prvků komparativního výzkumu, souvislosti reliability a validity výzkumu s ekvivalencí a komparabilitou zkoumaných jevů.

*Reliabilita* (spolehlivost měření) je jeden ze základních požadavků či vlastností měření ve výzkumu. Vyjadřuje přesnost ve smyslu stálosti výsledků měření a absenci chyb vznikajících při opakovaném měření téhož jevu za stejných podmínek. Z hlediska rozptylu naměřených hodnot tohoto jevu jde o minimalizaci náhodných chyb, které se při jeho měření objevují [Řehák, Řeháková 1986: 38; Hendl 2004: 48]. Kritérium reliability v případě komparativních šetření říká, do jaké míry jsou položky použité ve výzkumu mezi společnostmi neměnné a dostatečně obecné [Przeworski, Teune 1970: 11].

Pojem reliability je komplementární k pojmu *validity* (obsahová platnost, bezchybovost), která značí, do jaké míry si vzájemně odpovídá předem stanovený výzkumný cíl se skutečně dosaženými výzkumnými výsledky: *Změřili jsme skutečně to, co si myslíme, že jsme změřili?* Validita znamená důraz na splnění výzkumného záměru i získání věrného obrazu reality se všemi logickými vazbami. Pojem validity bývá ze statistického úhlu pohledu spojován se systematickou chybou, která se ve výzkumu vyskytuje [Řehák, Řeháková 1986: 38; Hendl 2004: 48]. Z hlediska mezinárodních výzkumů kritérium validity vypovídá o tom, do jaké míry měřicí nástroje aktuálně měří to, co je zamýšleno skutečně měřit, do jaké míry zobrazují položky pozorovaný fenomén tak, jak se objevuje v každém sociálním systému [Przeworski, Teune 1970: 11].

Celý koncept mezinárodní ekvivalence souvisí s možností jejího narušení. Rušivé faktory se v mezinárodních výzkumech seskupují do systematické chyby, která následně ohrožuje validitu měření. A s ohrožením otázky, zda měříme skutečně to, co chceme měřit v každé zkoumané společnosti, se automaticky pojí zpochybnění *komparability* měřených jevů v různých kulturních skupinách [Harkness, Mohler, Van de Vijver 2003: 13; Van de

Vijver 2003b: 207]. Takovéto determinující důsledky měření pro srovnatelnost jevů jsou na poli mezinárodních výzkumů vyjádřeny v konceptu *ekvivalence*, protože komparabilita a ekvivalence se pokládají za vzájemně související zrcadlové jevy [Van de Vijver 2003a: 144]. Pokud je díky hromadícím se chybám ve výzkumu narušena srovnatelnost zkoumaného jevu ve všech společnostech, je ohrožena celková ekvivalence mezinárodního výzkumu.

V ekvivalentním výzkumu musí dojít k naplnění tzv. konstruktové teoretické validity. Měření, na které se ve výzkumu soustředíme, musí ukazovat pozitivní korelaci s položkami, které měří identické konstrukty (convergent validity). Zároveň ale toto měření nesmí korelovat s žádnými jinými položkami, které měří jiné (neidentické) konstrukty (discriminant validity) [Groves 2004: 23]. Tyto dva jevy hovoří o naplnění teoretické validity, zmíněné korelace se testují v každé zemi a odkazují ke splnění ekvivalence daných měření.

V tomto kontextu je navýsost důležité vymežit často zmiňované pojmy indikátor, konstrukt a koncept, neboť spolu úzce souvisí. V obecném metodologickém schématu pojmy „*indikátor, konstrukt a koncept*“ hovoří o vztahu mezi teorií a empirií. Indikátory, které jsou v rozsáhlých mezinárodních výzkumech mnohočetné a rozvinuté a dají se přímo měřit, jsou zapotřebí k získání vhledu do postoje, který respondent zaujímá ke konstrukt, který přímo měřitelný není. Tento konstrukt pak dále představuje základní teoretický jev, který je předmětem měření, ale musí být z důvodu zjednodušení měření převeden do měřitelných (empirických) indikátorů, protože sám o sobě měřitelný není, je příliš složitý. Měřený indikátor ve výzkumu poskytuje spojení s latentním konstruktem, který zastupuje teoretický koncept [Harkness, Mohler, Van de Vijver 2003: 11]. Teoretickým konceptem se většinou myslí sociální fenomén, který je ve výzkumu zkoumán, který sociální badatele zajímá. Empirický indikátor je pak tento sociální fenomén jednoduše vyjádřený otázkou v dotazníku. Protože se komparativní výzkumy orientují na výskyt relativně složitých sociálních jevů, doporučuje se v dotazníku pro jejich popis použít mnohočetné indikátory, tedy větší množství otázek nebo různé baterie otázek. Ty by měly pokrýt pokud možno všechny dimenze sociálního jevu a dohromady vytvářet latentní konstrukt, který co nejlépe reprezentuje zkoumaný sociální jev.

I když pro zkoumané jevy (koncepty) používáme stejná pojmenování v různých společnostech, ještě to neznamená, že jsou totožné a srovnatelné. Významy jednotlivých pojmů se liší pro různé lidi, v různých dobách a v různých kulturách. Musíme hledat kritéria, podle kterých zhodnotíme, zda je srovnání konkrétních dvou jevů z různých společností vhodné. Komparabilita těchto jevů závisí primárně na jasném vymezení zkoumaného jevu a

porozumění jeho podstatě. Můžeme si vedle sebe postavit tři odlišné koncepty zainteresovaných skupin, zájmových skupin a nátlakových skupin, ve kterých se každá tato skupina liší množstvím komunikace, kterou aktivně vede s politickým vedením. I když se mohou zdát rozdíly mezi jednotlivými pojmy nepatrné, neporozumění teoretickému konceptu má v komparativních výzkumech vážné následky [Przeworski, Teune 1970: 10]. Některé sociální jevy dokonce ani nemají svůj vhodný ekvivalent v jiném sociálním systému a nejsou jednoduše srovnatelné.

Problémové nejsou v mezinárodním komparativním výzkumu jen obecné koncepty, ale i indikátory, které je mají ve výzkumu měřit. V komparativních šetření i malé rozdíly mezi indikátory zesilují významné rozdíly mezi zeměmi, proto musí být splněn alespoň základní princip pokrytí všech podstatných aspektů teoretického jevu použitými indikátory v každé kultuře. Navíc z důvodu naplnění validity výzkumu musí zvolené indikátory vyjádřit všechna specifika kulturního kontextu ve všech zkoumaných společnostech. Vezměme si příklad sociálního jevu, který se nazývá nepotismus, což je jistá forma protěžování příbuzných v práci nebo škole. Tento sociální fenomén je v západních vzorcích chování vnímán jako korupční jednání, na druhou stranu ve strukturách východního systému je stejné chování chápáno jako výraz loajality a odpovědnosti vůči vlastní rodině [Morris-Jones 1964]. Všechny tyto nuance musí měřicí indikátor zachytit, aniž by ale zcela zastínil obecné prvky zkoumaného konceptu. V komparativních výzkumech jde tedy především o budování mezinárodní validity operacionálních definic ekvivalentních teoretických pojmů.

## 4.2 Multikulturní ekvivalence

Komparativní výzkumy jsou na poli metodologie a měření jedním z nejkritizovanějších odvětví sociálněvědního výzkumu [Johnson 1998]. Pravidla pro jejich realizaci jsou z pohledu mnoha sociálních vědců daleko za vývojem metodologických pravidel národních výzkumů, a pokud dochází v rámci mezinárodních výzkumu k nekritickému přijetí metod a postupů, které bývají uplatňovány v národních výzkumech, dochází k ohrožení ekvivalence teoretických jevů i měření. Johnson [1998: 2] hovoří o tom, že komparativní výzkumy postrádají především metodologický konsensus pro stanovení *mezikulturní ekvivalence*. Tato absence konsensu podle něj vyplývá již z primárně špatného uchopení podstaty ekvivalence a nedostatečné klasifikace daného problému, které způsobují, že ekvivalence ve výzkumu není dosahováno jednotným systematickým způsobem.

Zásadní empirický problém v mezinárodních výzkumech, který má ale ve výsledku dalekosáhlé teoretické následky, je nalezení ekvivalentních fenoménů v různých prostředích a analyzování vztahů mezi nimi ekvivalentním (stejným) způsobem. Przeworski a Teune analyzovali na mezinárodním poli dva konvenční přístupy k těmto problémům měření [1966: 553].

Jeden přístup tvrdí, že mezispolečenské nebo interkulturní rozdíly jsou kvalitativní povahy. V každé společnosti existuje řada provázaných vnějších faktorů, které ovlivňují fungování každého jednotlivého sociálního jevu, a oddělení konkrétního jevu z jeho vlastního kulturního a sociálního prostředí není možné. Mezinárodní srovnávání tak nemůže být nikdy validní a ekvivalence nelze dosáhnout.

Druhý opačný přístup naopak prosazuje myšlenku, že mezinárodní validity jednotlivých identických či ekvivalentních položek popisujících různé kulturní jevy lze dosáhnout a že lze tyto položky porovnávat nezávisle na konkrétním kulturním prostředí. Praxe ve vědecké oblasti vypovídá o plošném přijetí druhého pohledu, neboť většina mezinárodních studií v sobě implicitně nebo explicitně zahrnuje srovnávání různých zemí a kultur na základě pozorovaných znaků.

Pokud si připustíme, že je možné mezinárodní validity skutečně ve výzkumu dosáhnout a že sociální jevy je možno porovnávat nezávisle na sociálním prostředí, problémy ekvivalence, které se s tím přirozeně pojí, jsou pak v mezinárodních výzkumech konfrontovány na dvou úrovních výzkumu. V první fázi na začátku výzkumu, kdy je nutné nastolení ekvivalentního měření jevu v různých zemích, a ve druhé fázi, pokud jsou mezinárodní data již analyzována a je nutné odstranit vliv vnějších proměnných [Przeworski, Teune 1966: 553].

#### **4.2.1 Ekvivalentní měření ve výzkumu**

Tak jako jakýkoli monokulturní sociálněvědní výzkum i mezinárodní výzkum požaduje určité výzkumné standardy. Avšak standardy jedné kultury nemusejí být aplikovatelné do jiné kultury [Przeworski, Teune 1966: 555]. Kupříkladu, může být intenzita sociální integrace lokálních komunit měřena stejnými indikátory v USA i Indii? Co konstituuje identické a co tvoří rovnocenné (ekvivalentní) měření stejných konceptů v různých zemích?

Przeworski a Teune [1966: 555] se kriticky staví proti používání jednoduchých stejných položek ve výzkumu k získání ekvivalentních, mezinárodně validních závěrů.

Totožnou povahu indikátorů mohou oslabit jazykové odlišnosti různých společností, časové rozvržení rozhovorů, které není ve všech zemích stejné, nesrovnatelná školení tazatelů atp. V hledání ekvivalence teoretických jevů se vědci musí dostat až za problém identických položek. Jednoduché stejné položky nejsou postačující k ekvivalentnímu měření konceptů na jakékoli úrovni analýzy [Przeworski, Teune 1966: 556; Jowell 1998]. Jak už bylo na začátku řečeno, jedná se o empirický problém, proto Przeworski a Teune nabádají badatele k pokusu přetransformovat význam jevů v různých kulturách do ekvivalentní podoby právě empirickým postupem (viz. Kapitola 4.3).

Ekvivalence ve výzkumu je možno dosáhnout také analýzou vztahů uvnitř konkrétních zemí a srovnáním těchto vztahů mezi všemi odpovídajícími systémy. Otázka pro komparativní přístup je, zda jsou vztahy mezi dvěma proměnnými stejné ve všech zemích nebo do jaké míry je tento vztah podobný. Pokud skutečně ve všech zemích vztahy proměnných fungují stejným ekvivalentním způsobem (ani jejich síla či intenzita se významně navzájem neliší), nebyly zpozorovány žádné systematické odlišnosti a zamítá se intervence třetí proměnné, která se většinou ve výzkumu odráží jako vliv kulturního nebo sociálního prostředí a znemožňuje nastolení ekvivalence ve výzkumu [Przeworski, Teune 1966: 566].

Van de Vijver [2003a: 145] upozorňuje na to, že nabývání validity a odstraňování systémových chyb nesmí být jen slepým empiricismem vědců. Pokusy o maximální ekvivalenci mohou někdy selhat a občas nedostatek či nemožnost ekvivalence ukazuje na přítomnost zajímavých mezikulturních rozdílů, jimž by měla být věnována stejná pozornost jako budování ekvivalence samotné.

#### **4.2.2 Kategorizace ekvivalence**

Metodologický konsensus pro stanovení ekvivalence ve výzkumu je neustále zpochybňován nejasnou kategorizací tohoto konceptu. Navíc mnoho sociálních vědců prosazuje jen jednu izolovanou techniku nastolování ekvivalence, která má poskytnout škálu pro reliabilní a validní měření stejných fenoménů v různých zemích, a zajistit tak jejich srovnatelnost.

Pojetí ekvivalence se vyznačuje existencí rozmanité klasifikace jak na obecné úrovni, tak i v oblasti konkrétních vymezení. Przeworski a Teune [1970: 126] určili na té nejobecnější úrovni *ekvivalenci slabou a silnou*, která má označovat míru spolehlivosti měření různorodých položek, jež má být přímo úměrná hodnocení ekvivalentních měření.

Čím spolehlivěji položky jevy měří, tím je ekvivalence jejich měření silnější. Johnson [1998: 6 - 7] naproti tomu identifikoval přes 50 různých konkrétních typů ekvivalence, ze kterých ale také vydělil dva obecné druhy.

První typ je tzv. *interpretativní ekvivalence*, která se primárně zabývá podobnostmi v pojetí sociálních jevů, jež jsou ve výzkumu nějakým způsobem interpretovány, a zdůrazňuje tak ekvivalenci významu pojmů. Pro Johnsona tento druh ekvivalence disponuje zcela zásadním významem. V konkrétním pojetí lze vymezit nespočet jejích příkladů: konceptuální (conceptual) ekvivalence, konstruktová (construct), kontextuální (context), funkcionální/funkční (functional) ekvivalence, lingvistická (linguistic), ekvivalence významů (meaning), teoretická ekvivalence (theoretical) atp. Naneštěstí interpretativní ekvivalence nemůže být stanovena pro kulturně specifické fenomény či koncepty, neboť nesdílejí stejné významy v různých kulturách [Johnson 1998: 11].

Druhý typ definuje tzv. *procedurální ekvivalenci*, která se naopak zaměřuje na měření a procedury mezinárodního srovnávání a vyzdvihuje konzistenci tohoto měření. Podle Johnsona nabytí procedurální ekvivalence není ve výzkumu vždy přímo nutné, ale je možné ji dosáhnout i pro sociální jevy ovlivněné kulturním prostředím. Svým charakterem do pojetí procedurální ekvivalence spadají ty skupiny ekvivalence, jež zajišťují komparabilitu mezi výběrovými šetřeními provedenými v různých kulturách a starají se o srovnatelnost výzkumných položek a škál. Jmenovitě to může být ekvivalence indikátorů (indicator), instrumentů měření (instrument), položek (item), ekvivalence měření (measurement), strukturální ekvivalence (structural), skalární (full score), faktorová ekvivalence (factor), přímá a relativní (direct/relative), operacionální (operational), ekvivalence ke kritériu (criterion), obsahová (content), psychologická (psychological) ekvivalence atp.

#### **4.3 Techniky nastolování mezinárodní ekvivalence**

Johnson [1998: 12] organizuje metody, jež mají za úkol být nápomocny v nastolování mezinárodní ekvivalence, podle čtyř specifických fází výzkumného projektu: vývoj otázek, pretest dotazníku, sběr dat a analýza dat. Ke každé této fázi předkládá několik bodů jako vymezení jednotlivých metod, které pro konkrétní časový okamžik výzkumného projektu představují nejlepší volbu vylepšení mezikulturní ekvivalence (viz. Tabulka 4.1).

##### *A) Fáze vývoje otázek*

V první fázi vývoje otázek se podle Johnsona [1998: 13] klade důraz na budování

interpretativní i procedurální ekvivalence. Interpretativní ekvivalence závisí na aktivní účasti, konzultaci a spolupráci národních expertů s hlubokou znalostí vlastní země, tímto způsobem mají vědci možnost reprezentovat svou kulturu. Vybudováním takovéhoho multikulturního a multilingvního týmu se výzkum šikovně vyhýbá hrožícím etnocentrickým tendencím. V této fázi výzkumného procesu vědci hodnotí vhodnost a relevanci specifických výzkumných položek, použití výzkumných nástrojů, metod sběru dat, přístup k překladům dotazníků a identifikují kulturně specifické vlivy, které se ve výzkumu mohou objevit. Často tento druh evaluace provádí s pomocí časově náročných kvalitativních metod, jejichž úkolem je v této souvislosti rozeznat kognitivní procesy, jež probíhají u různých osob v různých kulturách v průběhu dotazování.

Specifický přístup, který odpovídá myšlence interpretativní ekvivalence, je tzv. *facet analysis* [Johnson 1998: 20]. Facet analysis stejným způsobem zdůrazňuje sdílený význam pojmů v různých společnostech namísto „pouhé“ srovnatelnosti stimulů v dotazníku a mapuje výzkumné otázky a jejich dimenze. Může být používána i v procesu překladu dotazníků.<sup>5</sup>

Přístupu procedurální ekvivalence odpovídají návody na sestavení „dobré“ otázky a škály odpovědí. Rady se obecně týkají délky věty otázky (max. 16 slov), jednoduchosti a přímosti. Vymezují se obecně stylistická, syntaktická a gramatická doporučení pro tvorbu otázek do dotazníku nebo kategorií do škály odpovědí (např. vyhýbat se vágnosti, obecnosti, nejednoznačnosti, dbát na přeložitelnost, používat známé a často užívané výrazy, vyhýbat se slangu atp.) [Johnson 1998: 15].

#### *B) Pretest dotazníku*

Použití kognitivních rozhovorů v případě pretestu dotazníku směřuje podle Johnsona [1998: 20] k odhalení problémů, které respondenti mají s interpretací otázky nebo s porozuměním škále odpovědí. Z provedení kognitivních rozhovorů mohou vědci získat výstupy v podobě kódovaných chování zkoumaných respondentů. Ty pak odhalují interpretační problémy s konkrétní položkou nebo odpovědí a umožňují podle nich dotazník vhodně měnit a doplňovat. Sledováním alternativních způsobů sběru dat v pretestu dotazníků vědci navíc ještě pojišťují technickou ekvivalenci výzkumu.

#### *C) Sběr dat*

Pro měření konstrukt je v komparativním výzkumu výhodnější použít mnohočetné indikátory. Jak už bylo výše naznačeno, identické položky nemusí ve výzkumu zaručit rovnocennou validitu, proto je upřednostňováno použít pro jeden teoretický jev více různých

---

<sup>5</sup> Více o zpětném překladu a použití facet analysis Ingwer Borg [1998].



otázek a využívat různé odpovědní škály. Jako velice silný výzkumný nástroj se pak jeví kombinace univerzálních otázek, stejných ve všech zkoumaných zemích a kulturách (standardně označovaných jako *etic* otázek), a kulturně specifických otázek, přizpůsobených sociálnímu prostředí dané země a majících relativně unikátní význam (pojmenovávaných jako *emic* otázky). Terminologie *etic* a *emic* otázek je analogická k terminologii položek v dotazníku typických a stejných pro všechny kultury a položek specifických pro každou kulturu. Naplnění těchto principů může zajistit jak interpretativní ekvivalenci, tak i procedurální ekvivalenci.

Jako nutné se v otázce sběru dat jeví sladění respondenta s tazatelem, které má zajistit také procedurální ekvivalenci. Nejčastěji se využívá tzv. *matching procedure*, což je přístup, který upřednostňuje, aby si tazatel a respondent vzájemně odpovídali v sociodemografických charakteristikách, pohlaví, etniku, jazyku a aby byl tazatel „domorodec“ [Johnson 1998: 24]. Pokud se všechny tyto aspekty sladí, očekává se, že se bude respondent cítit uvolněněji a bude přístupnější k vyplňování dotazníku. Navíc někteří vědci předpokládají, že za těchto podmínek může také docházet k minimalizaci chyb pozorování, jež jsou způsobovány respondenty (*respondent error*). Pokud se v některých případech upřednostňuje výběr „odlišného“ tazatele, výzkumníci k tomu mají konkrétní důvody vyplývající z potřeb výzkumu.

#### *D) Analýza dat*

Všechny více méně psychometrické přístupy využitě v předešlých třech fázích výzkumného projektu jsou ve čtvrté fázi analýzy dat nahrazeny principy matematické statistiky. Analýza položek (resp. statistické třídění prvního a druhého stupně) mapuje variabilitu rozložení různých kategorií odpovědí v různých kulturních skupinách a korelaci jednotlivých položek v dotazníku. Ve své podstatě jednoduchá analýza položek statisticky testuje problematické jevy *response style*, sociální *desirability*, *reliability* škály atp (více Kapitola 5.2.2.).

Další již poněkud složitější statistické metody, např. konfirmační faktorová analýza nebo multidimenzionální škálování, shluková analýza atp., se orientují na procedurální ekvivalenci. Srovnávají vztahy mezi souborem položek měření v různých kulturních skupinách, případně nahrazují příliš souhrnné indikátory kultury (*rasa*, *etnikum*, *země původu* atp.) více specifickými položkami, které už představují odlišné rysy zkoumaných společností.

Jako sofistikovanější přístup k identifikaci fungování výzkumné otázky v různých kulturách a k rozpoznání, do jaké míry odpovídá daná otázka zkoumanému jevu, se jeví *Item*

Response Theory (IRT). Výstupy Item Response Theory nejvíce přispívají k dosažení ekvivalence překladu položek a ke zjištění, zda správně odpovídají kategorie na škále odpovědí potřebě výzkumníka a zda jsou škály kompletní [více Borg 1998; Van de Vijver 2003b: 216]. Teorie zobecnitelnosti (Generalizability Theory) pracuje s chybovými složkami měření, které sleduje analýzou rozptylu. Zaměřuje se speciálně na různorodost odpovědí způsobenou tím, že v každé zemi se v dotazníku používá odlišný jazyk. Chyby, které z použití různých jazyků vyplývají, testuje na bilingvních respondentech, kteří mají vypracovat dotazníky v různých jazycích a od kterých se očekává, že budou odpovídat stejným způsobem ve všech jazycích [Johnson 1998: 26]. K testování ekvivalence měřících položek v procesu analýzy dat lze také využít i různé jiné, teprve se rozvíjející moderní techniky, například tzv. SEM techniky (Structural Equation Modeling) [více Billiet 2003].

Významná a velice podnětná strategie nastolování ekvivalence v analýze dat je tzv. „*identity-equivalence procedure*“, kterou blíže rozpracovávají Przeworski a Teune [1966, 1970]. O proceduře „*identity-equivalence*“ autoři tvrdí, že umožňuje výzkumníkovi vzít v úvahu sociální jev, který by se jinak jevil kvalitativně odlišný v různých kulturách a z tohoto důvodu neměřitelný. Vnitřní filozofie tohoto přístupu staví na rozdělení celkového souboru indikátorů, které měří sociální fenomén v každé kultuře, na dva podsoubory položek:

- ◆ *položky typické pro všechny kultury* (etic otázky) – tento soubor v sobě zahrnuje všechny položky v dotazníku, které jsou pro všechny kultury stejné, pokud je tento soubor společných položek pro všechny kultury homogenní (ukazuje stejnou vlastnost), jeho položky jsou definované jako identické (v mezinárodní analýze mají identickou mezinárodní validitu)
- ◆ *položky specifické pro každou zemi* (emic otázky) – tento podsoubor položek je analyzován odděleně pro každou zemi, položky v něm obsažené jsou odlišné v každé zemi a vyznačují se obtížným překladem v různých kulturách; specifické položky z tohoto podsouboru, které jsou v určitých zemích zkorelované s identickými položkami z předchozího souboru společných položek, jsou vybrány jako ekvivalentní (mají ekvivalentní mezinárodní validitu)

Pokud používáme tuto metodu, není nutné v praxi začínat s podsouborem specifických položek pro různé země. Je možné ze začátku výzkumu postupovat s dostatečně velkým počtem indikátorů použitých ve všech zemích a oddělit identické a ekvivalentní položky až v procesu statistické analýzy. Specifické položky se v analýze dat vyznačují velkou variancí mezi zeměmi a nebo nejsou zkorelovány s ostatními položkami v dotazníku ve více než jedné

zemi [Przeworski, Teune 1966: 559].<sup>6</sup>

Tabulka 4.1. Metody získávání ekvivalence v mezinárodních výzkumech

<b>A.</b>	<b>Fáze vývoje otázek</b>
	1. Expertní konzultace/spolupráce
	2. Etnografický (či jiný) kvalitativní přístup
	3. Konstrukce „dobré“ otázky
	4. Relevantní překlad
	5. Facet analysis
<b>B.</b>	<b>Pretest dotazníku</b>
	6. Kognitivní rozhovory/strukturální sonda
	7. Měření frekvence odpovědních kategorií
	8. Kódování chování respondenta
	9. Komparace alternativních způsobů sběru dat
<b>C.</b>	<b>Sběr dat</b>
	10. Užití vícedimenzionálních indikátorů
	11. Užití univerzálních (etic) i kulturně-specifických (emic) otázek
	12. Sladění respondenta a tazatele
<b>D.</b>	<b>Analýza dat</b>
	13. Analýza položek
	14. Item Response Theory
	15. Generalizability Theory
	16. Konfirmační faktorová analýza
	17. Multidimenzionální škálování
	18. Kontrolní (statistické) proměnné
	19. Metoda „identity-equivalence“

Zdroj: převzato z [Johnson 1998: 12]

Jak už bylo výše řečeno, nejedna odborná publikace věnující se komparativním mezinárodním výzkumům se zaměřuje na vydělení pouze jedné konkrétní metody (většinou jako originální příspěvek sociálního vědce v dané studii), která by jako jediná měla být zodpovědná za možnost srovnání různých proměnných a pokrytí mezikulturní ekvivalence

<sup>6</sup> Zajímavé srovnávání se také v analýze dat provádí porovnáváním variance identických položek s variancí ekvivalentních položek pro každou kulturu, tento index UR vypočítá „universality of relevance“ (čím menší je, tím více položek je specifických pro danou zemi) [Przeworski, Teune 1966: 568].

ve výzkumu. Tento pohled je značně omezený, neboť z perspektivy snahy o dosažení kvalitního mezinárodního šetření je nezbytné ekvivalenci vnímat multidimezionálně a zdaleka přesáhnout hranice jedné jediné ideální metody.

Jak bude řečeno v další kapitole, kvalitu komparativních výzkumů lze zajistit synchronizovaným využitím tří principiálně odlišných přístupů, které jsou ale vzájemně propojeny: technickým přístupem, matematickou statistikou a psychometrickými přístupy. Johnson [1998] navíc seřazuje metody nastolování mezinárodní ekvivalence podle čtyř specifických fází výzkumného projektu: vývoj otázek, pretest dotazníku, sběr dat a analýza dat. V prvních třech fázích hrají prim psychometrické přístupy, zatímco ve čtvrté fázi analýzy dat nachází své uplatnění především matematická statistika. Nejenže mezinárodní ekvivalence je ve spojení s celkovou kvalitou výzkumů dosahováno dvěma odlišnými způsoby, k jejímu potenciálnímu selhání se v různých úsecích realizovaného výzkumného projektu přistupuje z různých hledisek. Není vhodné z tohoto důvodu zanedbávat širší kontext v dosahování mezinárodní ekvivalence a kvality mezikulturního srovnávání a izolovat nebo upřednostňovat pouhou jednu techniku nebo jeden přístup.

## 5 Dosahování kvality v mezinárodních komparativních šetření

Mezinárodní komparace je specifická velkou organizační, finanční i časovou náročností pořizování kvantitativních dat, obzvláště pokud jsou kladeny nároky na kvalitu provedení výzkumu. Požadavek efektivity výzkumu v této souvislosti dostává zásadní význam. Mezinárodní šetření, shromažďování a harmonizace statistických údajů nebo kompilace mezinárodní databáze z jiných zdrojů proto obvykle nejsou realizovány pro účely jednoho projektu, ale vznikají s cílem vytvořit zázemí pro širší výzkum vhodný pro potřeby většího množství projektů [Krejčí 2006: 149 – 150].

Mnohé studie poukazují na fakt, že i přes zřejmou náročnost rozsáhlých komparativních výzkumů mají sociální vědci sklon akceptovat v mezinárodních šetřeních méně přísné standardy než v národních či monokulturních studiích. Jowell [1998] odhaluje možnou příčinu v přání vědců respektovat odlišné kulturní normy, ze kterého však plyne nevhodné tolerování nesrovnatelných výzkumných metod. Tyto problémy se ale neobjevují jen v perspektivě komparativních studií založených na národních zdrojích dat, které nejsou v první řadě navrhovány ke srovnávání zemí, ale také v nejpečlivějších komparativních projektech, speciálně navržených k získání ekvivalentních mezinárodních měření.

Tyto problémy mezinárodních výzkumů jsou neustále prezentovány v určité podobě i v národních výzkumech, jelikož žádný národ není homogenní, pokud jde o slovní zásobu obyvatel, způsob vyjádřování, stupeň vzdělání atp. Kvantitativní národní výzkum závisí z hlediska jeho reliability na zásadě rovnocennosti. Aby se splnily požadavky reprezentativity, pravděpodobnosti výběru jednotek v národním vzorku by měly být rovnocenné (nebo přinejmenším známé a nenulové), stejně tak musí být různé subskupiny společnosti rovnocenně zastoupené. Otázky musí mít též široce ekvivalentní význam pro všechny respondenty, kvůli ujištění, že variance mezi daty pochází z rozdílů mezi jejich odpověďmi na otázky v dotazníku, spíše než z rozdílů v interpretaci samotných otázek. Pouze když dosáhnou národní výzkumy úspěchů v těchto cílech, tak se jejich výstupy pravděpodobněji blíží sociální realitě. I když je heterogenita stále prezentována jako problém výzkumů na národní úrovni, je to hlavní překážka úspěchu mezinárodních výzkumů [Jowell 1998].

Výzkumníci se při realizaci mezinárodních šetření musí vyrovnávat se specifickými rozdílnostmi daných zemí vybraných do mezinárodního výzkumu a jejich metodologickými a výzkumnými zvyklostmi, které se týkají preferovaných způsobů dotazování, hluboce zakořeněných typických výběrových metod, rozdílů v uznávání přijatelnosti návratnosti dotazníků, odlišného používání vizuálních pomůcek při dotazování, rozdílných metod

v proškolení tazatelů, nestejně sociodemografické klasifikace atp. [Jowell 1998]. Na druhou stranu reakce výzkumníků na tyto komplikace v podobě přílišné tolerance méně přísných a nesrovnatelných metod jistě není adekvátní. Mohly by být totiž vážným zdrojem omylů a systematických chyb vedoucích v analýze k radikálním závěrům o rozdílnostech mezi národy, a přitom by byly tyto odlišnosti zemí způsobené jen užitím rozdílných metod v různých kulturních prostředích.

Důležitost a užitečnost přesných a pečlivých měření mezinárodních sociálních výzkumů je neoddiskutovatelná. Herkenrath [2002: 517] navíc dodává, že základní otázka výzkumné preciznosti a přísných pravidel nabývá jak vědecké, tak i politické důležitosti, pokud mají výsledná zjištění vědců očekávatelný vliv na politickou praxi. V tuto chvíli výsledky měření založené na mylných analýzách nemusí být pouze nesolidní a nespolehlivé, ale i potenciálně devastující v aplikaci do sociálního světa.<sup>7</sup>

V boji proti přetrvávajícím pochybám badatelů ohledně faktické srovnatelnosti, validity a významu mezinárodních výzkumů se budují rozsáhlé výzkumné programy (např. ISSP, ESS atp.) usilující o skutečně přísné technické standardy, za které se musí ve výzkumu zaručit každý potenciální člen projektu. Na místo klasického „*safari-modelu*“ mezinárodních výzkumů, který hovoří o tendenci výzkumného týmu vědců z rozvinutých zemí vyvíjet výzkumný plán a uplatňovat měřicí nástroje na jiné společnosti teprve se rozvíjejících zemí [Smith 2003: 86; Hantrais, Mangen 2006], operují současné mezinárodní projekty spíše s určitým typem demokratického modelu s mezinárodním přístupem a multikulturním týmem sociálních vědců [Jowell 1998].

Vzhledem k důležitosti komparativních studií se oblast kvality výzkumu neustále rozvíjí, i když se ještě stále považuje za nejokřídlenější překážku realizace mezinárodního srovnávacího výzkumu. Krejčí [2008: 19] tuto orientaci na kvalitu výzkumu správně posoudil jako pozornost věnovanou řešení mnoha důležitých problémů, které teprve mohou nastat. Logika kvality výzkumu se potom přirozeně zakládá na nezbytných kompromisech vztažených k celkovému výsledku, není jen otázkou například velikosti vzorku a tématického okruhu. Przeworski a Teune [1966] v kontextu metodologických problémů spojených s kvalitou výzkumu vidí jako řešení neustálé testování komparativní validity opakovaným prováděním mezinárodních výzkumů, které může sloužit k postupnému odhalení výzkumných komplikací, jež by bylo možno později zahrnout do výzkumného designu. Tento postup se může osvědčit i na teoretickém poli, protože opakovaným testováním stejných teorií

---

<sup>7</sup> Herkenrath (2002) uvádí příklady z výzkumů v rozvojových zemích.

v různých sociálních prostředích lze dosáhnout kvalitnějších obecných výroků, které jsou také v centru pozornosti komparativních výzkumů.

Pro souběžné vyhodnocení validity a spolehlivosti měření a dosahování celkové kvality komparativních šetření se dají v sociálních vědách identifikovat tři různá pojetí, jež vyžadují specifické přístupy jak k plánu výzkumu, tak i k analytickému zpracování získaných dat. Prvním významným přístupem v dosahování kvality šetření je *statistický přístup*, který se ve výzkumech objevuje v podobě vybudovaných technických standardů, jež by měly všechny země účastníci se mezinárodního programu dodržovat. Vycházejí z něj především principy výběrového designu a výpočty chyb měření. Druhým podobným pojetím je *technický a administrativní přístup*, který se věnuje budování pravidel ohledně technik sběru dat a vytváření standardní dokumentace výzkumu.

Závěry o validitě výzkumu ale nemohou být založeny pouze na statistických metodách nebo administrativních principech, proto se doporučuje kombinovanat strategie kvantitativních a kvalitativních (interpretativních) metod. Z tohoto důvodu jsou třetím významným zdrojem v dosahování kvality komparativního výzkumu *psychometrické přístupy*, které se soustředí především na konstrukci kvalitního nástroje měření.

## **5.1 Statistický přístup k dosahování kvality šetření**

### **5.1.1 Výběrový design**

Teorie a praxe výběru vzorku pro multipopulační studie ještě stále zůstaly nepříliš rozvinutým aspektem sociálního výzkumu. Jednak z důvodu nekvalitní dokumentace již uskutečněných srovnávacích šetření, jednak kvůli omezením jejich náležitého provedení na mezinárodním poli [Häder, Gabler 2003: 117]. Problémy vyvstávající v této metodologické oblasti jsou kladeny za vinu nekonzistentní teorie výběru vzorků, nevhodných aktuálně používaných výběrových designů a také jsou dílem časté kombinace náhodných a kvótních výběrů v rámci jednoho mezinárodního projektu. Kish [1994: 173] v otázce výběrového designu argumentuje tím, že „(v)ýběrové designy mohou být vybrány flexibilně a neexistuje zde žádná potřeba jejich podobnosti. Flexibilita výběru je obzvláště vhodná pro mezinárodní srovnávání, protože výběrové prostředky skvěle rozlišují mezi zeměmi.“

Je nutné si ale podle Kish [1994] uvědomit, že i tato flexibilita v mezinárodních srovnávaních předpokládá striktně pravděpodobnostní výběrovou metodu (resp. znalost

pravděpodobnosti výskytu každé populační jednotky) z důvodu následujícího usuzování na celou populaci z výsledků měření. Kvalita a reprezentativita výběru je tedy nezbytným předpokladem pro komparativní výzkumy.

Optimální design výběru vzorku sestává z naplnění několika požadavků. Výzkumníci musí vyřešit především výběrový rámec, to znamená nalézt oporu výběru, zdůvodnit případné nahodnocení či podhodnocení výběru, sepsat vhodnou definici populace, kterou se budou při výběru vzorku řídit všechny země účastníci se mezinárodního projektu. Příkladem definice populace je například jasné vymezení dospělé osoby, kterou lze do vzorku zahrnout, obzvláště upřesnění její dolní a horní věkové hranice, neboť toto pojetí se v tradicích výzkumu v různých zemích liší.

Dalším požadavkem je zvolení konkrétního druhu náhodného výběru. Výzkumníci standardně nevolí prostý náhodný výběr a naopak upřednostňují možnosti víceetapového náhodného výběru nebo stratifikovaného výběru. Systematický výběr podle opory se střetává s problémem nalézení vhodné opory ve všech participujících zemích, výběr skupin prvků také není příliš častý. Kvótní výběry rozsáhlé mezinárodní komparativní projekty zásadně nevyužívají kvůli jeho kvazi-reprezentativnosti a neprůhledné metodice, zejména při mapování návratnosti dotazníků.

Tím, že každá země má většinou možnost volby konkrétního druhu náhodného výběru podle vlastních výzkumných možností, výzkumníci mají za úkol zjišťovat designový efekt pro všechny země. Designový efekt reaguje na výběrovou chybu, která touto odlišnou volbou náhodných výběrů vzniká, a vyjadřuje tak svou povahou způsob srovnání a hodnocení různých výběrových designů realizovaných v různých zemích. Vypočítává se tak, že je realizovaná strategie zvoleného náhodného výběru srovnávána se standardní strategií prostého náhodného výběru, který (ve výpočtu) používá vzorek o stejné velikosti jako je vzorek vybraného náhodného výběru [Häder, Gabler 2003: 122; Krejčí 2008: 63].

V souvislosti s designovým efektem se musí vědci ještě rozhodnout o použití vážení, které má funkci korekce specifických rysů designu výzkumu a má tak za úkol redukovat systematickou chybu ve výzkumu, která je specifickými designy způsobována. Toto vážení se nazývá výběrová váha a je považována za opak principu stejné pravděpodobnosti výběrových jednotek a principu rovnocennosti, neboť tuto pravděpodobnost upravuje.

Posledním požadavkem pro sestavení optimálního designu výběru vzorku je výběr okruhu zemí do studie, ve kterých budou sociální jevy sledovány. Výběr zemí ke komparaci musí být nějakým způsobem racionálně zdůvodněn, neboť vhodným výběrem lze podpořit



i nepravdivou tezi.

Výběr zemí je jen zřídka náhodný, protože náhodný výběr zemí produkuje velké množství jednotek do výzkumu a není ve své podstatě reálně možný [více Van de Vijver, Leung 1997]. Mezinárodní výzkumy mají spíše kvazi-experimentální podobu a jsou realizovány systematickým výběrem zemí na základě teoretických či historických předpokladů jejich příbuznosti a podobného vývoje, aby nebylo srovnáváno nesrovnatelné [Häder, Gabler 2003: 122]. Taktická volba zemí je řízena principem nejlepší možné kombinace zemí zvláště pro potřeby výzkumníků. Například do výběru podobných zemí se mohou někdy zařadit i tzv. referenční případy, čímž výzkumníci podtrhují pozitivní líčení výskytu určitého sociálního jevu vkládáním negativních příkladů ze zemí, ve kterých se daný sociální jev nevyskytuje či vyskytuje v jiné formě.

Z hlediska výběru zemí ke komparaci existuje několik přístupů. Jeden přístup primárně zdůrazňuje ve výběru větší množství zemí ke srovnávání, na druhou stranu ale vznáší požadavek na podobnost jejich kultur, politických systémů, stupně hospodářského rozvoje, geografické polohy, sociální struktury atp. Tento přístup svou povahou odpovídá výběrové strategii, kterou Przeworski a Teune [1970: 32] vymezili jako maximální strategii (*Most Similar Systems Design*). Jedná se o techniku výběru těch nejpodobnějších zemí, které tvoří pro komparativní výzkum optimální vzorek. Nemusí se jednat jen o země, výzkumníci si mohou volit podobné regiony, vesnice, kmeny atp. Výhoda této strategie spočívá v tom, že pokud nalezneme nějaké rozdíly mezi podobnými zeměmi, počet proměnných vysvětlující tyto odlišnosti je dostatečně malý, menší než u vzdálenějších, méně podobných zemích. Klasicky se aplikuje ve studiu politických systémů, volebního chování, sociální mobility, sebevraždy atp.

Další variantou je podřízení výběru zemí cíli komparace, kdy je nutné si nejprve jasně definovat cíl výzkumu a následně vybrat reprezentativní země k dosažení tohoto cíle. V tomto směru je vhodné zvolit země, které se co nejvíce odlišují ve sledovaném znaku, ale v ostatních znacích jsou si co nejvíce podobné. Přístup odpovídá druhé výběrové strategii, kterou Przeworski a Teune [1970: 35] definovali jako *Most Different Systems Design*. V analýze dat se postupuje následující logikou, pokud se skupiny populace z různých zemí v pozorovaném jevu neliší, rozdíly mezi společnostmi se již dále nezvažují. Pokud bude například míra sebevraždy stejná v Rusku, Švédsku i Anglii, všechny proměnné, které odlišovaly tyto tři společnosti, jsou nadále v analýze dat irelevantní pro vysvětlení sebevraždy.

### 5.1.2 Chyby měření v komparativním výzkumu

Chyby měření, které se v mezinárodním výzkumu běžně vyskytují, jsou definovány jako odchylky měření ovlivňující empirické výsledky tak, že rozdílnosti nebo podobnosti mezi zeměmi jsou jen něčím uměle stvořeným při analýze dat [Braun 2003b: 137]. Krejčí [2008: 25] zdůrazňuje, že pro koncept kvality šetření je přínosné definovat celkovou chybu měření a vydělit její různé dimenze, ale pokud ji chtějí vědci ve výzkumu minimalizovat, je opět nutné zvážit chybu měření jako celek.

Termín celkové chyby měření (total survey error) klasicky zahrnuje kombinaci dvou prvků zkreslení. Jeden typ chyby měření je *systematické zkreslení* (bias), druhým prvkem celkové chyby jsou *nesystematické chyby* (náhodná odchylka – variance), obě části celkové chyby měření mají za následek parciální nebo nesprávné závěry studie. Systematická chyba je konstantní chyba, která se jednoduše definuje jako odchylka měření od pravdivé hodnoty a objevuje se při každé aplikaci daného designu výzkumu. Nesystematická chyba vypovídá o tom, že každé měření má svůj rozptyl (hodnoty jsou různé u jednotlivých jednotek výzkumu), tato chyba se tedy liší u jednotlivých aplikací designu výzkumu. Groves [2004: 10] vedle konstantní chyby (bias) a chyby rozptylu měření (variance) rozlišuje další úroveň klasifikace jednotlivých zdrojů chyb.

První Grovesova podúroveň vymezuje chyby pozorování a chyby chybějících pozorování. Obě mají jak náhodnou, tak i systematickou povahu. *Chyby chybějících pozorování* se vyznačují neuskutečněným měřením u části populace, jedná se o výběrovou chybu (sampling error), chybu pokrytí (coverage error) a chybu výpadků návratnosti (nonresponse error), z nichž každá má jiné důsledky pro povahu dat, a je tím pádem také různě závažná s ohledem na různé analytické záměry [Vinopal 2008: 13]. *Chyby pozorování* jsou definovány jako rozdíly odpovědí respondentů od pravdivých hodnot, kromě respondentů (respondent error) jsou jejich zdrojem tazatelé (interviewer error), měřicí nástroje (instrument error) a metody sběru dat (mode error).

Daleko detailněji se v kontextu mezinárodních měření rozpracovává systematická chyba, do které se koncentruje skupina rušivých faktorů, které vyplývají ze strukturálních a kulturních rozdílů mezi zeměmi. Tyto rušivé faktory pak zpochybňují komparabilitu měření v různých kulturních prostředích [Harkness, Mohler, Van de Vijver 2003: 13]. S narušením srovnatelnosti měření se pojí i zasažení ekvivalence výzkumu, neboť srovnatelnost i ekvivalence vyrůstají z aplikace měřicího nástroje v nejméně dvou kulturních skupinách a následujícího srovnávání výsledků měření [Van de Vijver 2003a: 145].

Van de Vijver a Leung [1997] identifikovali tři typy systematické chyby a současně s tím i strategie, jak se s nimi vypořádat:<sup>8</sup>

- ◆ *konstruktová systematická chyba* (construct bias) – konstruktová systematická chyba odkazuje k problémům s neidentickými a nesrovnatelnými teoretickými jevy mezi kulturami, znemožňuje použít identické položky v dotazníku, protože ty nepokryjí všechny podstatné aspekty zkoumaného jevu
- ◆ *metodologická systematická chyba* (method bias) – metodologická systematická chyba se týká všech zdrojů chyb vyrůstajících z metodologických aspektů výzkumu (nekomparability vzorků – sample bias; rozdílů položek v dotaznících – instrument bias; vlivu tazatelů a respondentů a jejich komunikace, rozdílné administrace dotazníků – administration bias atp)
- ◆ *položková systematická chyba* (item bias) – položková systematická chyba popisuje nejružnější odchylky na úrovni položek v dotazníku, s tím souvisí jejich odlišné překlady, rušivé faktory a kulturní specifika, které položkám nedovolí, aby se v dotazníku používaly identicky ve stejném znění, protože musí být pro kulturní prostředí vhodné

Čím dál větší pozornost je v komparativních výzkumech také věnována *otázce návratnosti* ovlivňující kvalitu a spolehlivost sebraných dat [Couper, De Leeuw 2003: 157]. Některé skupiny obyvatel je velice obtížné zastihnout a nebo dotazovat, kupříkladu nadmíru zaměstnané respondenty, respondenty v některých skupinách povolání, subskupiny obyvatel se specifickým životním stylem, skupiny obyvatel přebývajících v nepřístupných institucích (vězení, psychiatrické léčebny atp.), sociálně slabé skupiny, exkludované skupiny atp. Ve výzkumu může docházet i k tzv. podvýběru důležitých skupin, příkladem je zjišťování sociální mobility, u kterého se vyskytuje obrovský problém sehnat respondenty, kteří takto mobilní skutečně jsou, protože jednoduše nejsou doma.

Ukazatel, který sleduje podíl vydaných dotazníků k dotazování (na platných adresách výběru) a těch, které jsou skutečně sebrány, vyplněny a zpracovány, se nazývá návratnost, resp. míra návratnosti (*response rate*) [Krejčí 2008: 73]. Chybu, která se v rámci návratnosti vyskytuje a souvisí s nenáhodnými rozdíly mezi respondenty a nerespondenty, můžeme označit jako chybu výpadků návratnosti (*non-response error*), a znamená selhání realizovat měření na celém vybraném vzorku [Krejčí 2008: 72].

Výpadky návratnosti bývají sledovány jednak u jednotlivých položek v dotazníku

---

<sup>8</sup> více Van de Vijver [2003a: 145 – 149; 2003b: 207 – 233]

(item nonresponse), kdy někteří respondenti neodpoví na konkrétní otázky v dotazníku, a jednak u jednotek výběru (unit nonresponse), zahrnující jednotky z vybraného vzorku, u nichž dotazování z různých důvodů vůbec neproběhlo [Krejčí 2008: 72]. Respondenti v podstatě rozhovor odmítli nebo přerušili, nebo nebyli reálně kontaktováni z různých důvodů; zemřeli, byli jazykově nezpůsobilí, či psychicky nebo mentálně neschopni rozhovoru [Krejčí 2008: 73].<sup>9</sup> Komparativní výzkumy se spíše zaměřují na druhý typ výpadku návratnosti, na výpadek celé jednotky z výběru.

Zdroje příčin znemožnění kontaktu s vybranou jednotkou se mezi zeměmi díky kulturním odlišnostem značně liší. Příčiny výpadku návratnosti dotazníků se liší i mezi jednotlivými výzkumy díky odlišným výběrovým designům a díky obecně vnímané hodnotě sociálních výzkumů ve společnosti, toto vnímání výzkumů společností určuje do jaké míry se budou respondenti ochotni výzkumu účastnit [Couper, De Leeuw 2003: 165]. Velice často hraje důležitou roli tzv. administrativní, organizační příčina výpadků (např. selhání tazatelů). V některých agenturách tvoří až třetinu všech výpadků právě selhání tazatelů, tazatelé nevyřešili odmítnutí rozhovoru, lišili se mezi sebou mírou spolupráce, udělali chybu při zpracování atp. [Krejčí 2008: 86].

K řešení nonresponse, pokud nemůžeme chybějící údaje z analýzy vyloučit, nabízí statistika dva základní přístupy: imputaci a vážení [Krejčí 2008: 82]. Tato náhrada jednotek (resp. úprava po skončení sběru dat) v mezinárodním komparativním výzkumu však téměř nepřípadá v úvahu, neboť podle Coupera a De Leeuw [2003] deskriptivní statistiky, které numericky nahrazují chybějící údaje, ovlivňují vztahy mezi proměnnými a zkreslují výslednou analýzu dat. Jowell [1998] také poznamenává, že byly v praxi zachycené případy nahrazování nekontaktovaných respondentů například kvótní metodou, která způsobila deformaci pravděpodobnostního výběru a snížila hodnoty systematických chyb výběru vzorku.

Z těchto důvodů se preferuje řada metod používaných ke zlepšení návratnosti již při přípravě šetření a v průběhu sběru dat, u kterých Krejčí [2008: 82] ale upozorňuje na to, že i takto zvolená metodika (preventivního) zvyšování návratnosti může ovlivňovat velikost chyby návratnosti. Proto se v mezinárodních šetřeních klade velký důraz na pečlivou dokumentaci míry návratnosti a podrobných charakteristik nerespondentů, neboť díky ní

---

<sup>9</sup> Tzv. nevhodné výběrové jednotky: jedinec zesnul, není mu 15 let, komplex budov či byt je v demolicí (neobyvatelný), respondent emigroval, neobyvatelná adresa – školy, kanceláře, instituce, továrny, částečně obývané domy (letní sídla, chalupy), nepřístupné adresy, respondent není přítomen, je nemocný, nehovoří jazykem výzkumu atp.

výzkumníci lépe porozumí odlišnostem respondentů a nerespondentů, vyberou vhodnou metodu pro zvýšení návratnosti a získají empirický základ pro případnou statistickou úpravu [Couper, De Leeuw 2003: 177].

## **5.2 Technicko-administrativní přístup ke kvalitě výzkumu**

### **5.2.1 Metody sběru dat**

Výběr konkrétního způsobu sběru dat v komparativních mezinárodních výzkumech závisí na mnoha faktorech, například na rozpočtu výzkumu, harmonogramu, výběrové populaci, výběrovém designu, výzkumném tématu, komplexitě dotazníku atp. Klasicky se využívá ve volbě mezi odlišnými styly sběru dat možností standardizovaného dotazování rozhovorem (face-to-face interview), telefonního rozhovoru a elektronického počítačového výzkumu (CATI computer-assisted telephone interviews/CAPI computer-assisted personal interviews) [Skjåk, Harkness 2003: 179].

Zvolené metody sběru dat jsou v komparativních šetřeních vedle tématu otázek, stylu otázek a typu odpovědních škál příčinnou odlišností v odpovědích respondentů a zdrojem narušení vztahů mezi proměnnými [Saris 2003a: 268]. Tyto odlišnosti v odpovědích respondentů mohou být za prvé sociokulturní povahy. Na respondenty vždy určitým způsobem působí styl komunikace s tazatelem, který je jednoznačně determinován zvolenou metodou sběru dat. Realizovaná technika sběru dat určuje například pozici kontroly mezi tazatelem a respondentem a nebo se různým způsobem vyrovnává s výskytem osobních otázek, resp. odpovědí. V případě příliš osobních dotazů se někdy doporučuje využít samovyplňovacího dotazníku, někdy je raději zvolen rozhovor s promyšleným přístupem vyškolených tazatelů, tyto volby závisí většinou na konečném zvážení výzkumníka. Sociálně desirabilní odpovědi, ve kterých se respondenti vyhýbají reakcím považovaným za společensky nežádoucí, se také vyskytují v různé míře v závislosti na vybrané metodě sběru dat. Za druhé mohou být odlišnosti v odpovědích respondentů faktory kognitivních procesů, které probíhají v myslích respondentů v průběhu interpretace položené otázky a formování vhodné odpovědi do dotazníku [Skjåk, Harkness 2003: 181].

Pro komparativní výzkumy existuje způsob optimalizace výběru vhodného přístupu ke sběru dat. Tato optimalizace spočívá v nastolení konkrétní strategie volby metody sběru dat jako technického standardu, který bude plošně dodržován ve všech zemích účastnících se

mezinárodního projektu. Pro výběr dané metody se uplatňuje na jedné straně „*unimode design*“, který preferuje realizovat ve všech zemích stejný způsob sběru dat; tato konzervativní strategie má název „*keep-the-same-mode strategy*“. Tento přístup uplatňuje jednoduchý postup překladu otázek v dotaznících v rámci jedné stejné metody sběru dat. Unimode design ale trpí nízkou návratností, chybou pokrytí a někdy i systematickou chybou, záleží na kulturních zvláštностech a výzkumných tradicích daných zemí.

Na druhé straně se jako flexibilnější varianta sběru dat jeví „*mix-mode design*“ [Dillman 2000]. Hlavní myšlenka tohoto přístupu spočívá v tom, že pokud bude design výzkumu navržen obezřetně, tak i data sebraná různými metodami v různých zemích mohou být srovnatelná (kombinace osobních rozhovorů, telefonických rozhovorů, počítačového sběru dat atp.). Zvláštní důraz je kladen na možnost výběrů vhodné metody sběru dat podle preferencí a výzkumných tradic jednotlivých zemí (popř. finančních důvodů). Tento způsob sice disponuje vyšší návratností a lepším výběrem vzorku respondentů, protože může adekvátně vystihnout lokální podmínky, problematický ale může být nekonzistentní efekt metod sběru dat [Skjåk, Harkness 2003: 192].

Zmíněný nekonzistentní efekt metod sběru dat spočívá v tom, že položky v dotaznících interagují s konkrétní metodou (sběru dat), kterou jsou měřeny. Z této interakce položek a zvolené techniky vyplývá vznik náhodných i systematických chyb ve výzkumu. Klasické nástroje měření reliability a validity, kdy je jedna proměnná měřena jednou metodou, neumožňují odlišit, jaká část variance je náhodnou chybou měření a jaká je systematickou chybou metody [Vinopal 2008: 60]. Z tohoto důvodu se začal využívat tzv. *multitrait-multimethod design* (MTMM). MTMM je metodologický přístup, který chce vysvětlit rozdíly v korelacích mezi položkami a odlišnými metodami sběru dat, které je měří, proto provádí opakované měření několika proměnných různými metodami současně. Takto lze efekt zvolené metody nejlépe zachytit. Většinou využívá pro své analýzy kombinaci minimálně tří položek v dotazníku a tří využitých metod. Poprvé byl MTMM design představen Campbellem a Fiskem v roce 1959, kteří zjistili, že reliability, validita a efekt zvolené metody se liší mezi všemi metodami sběru dat [podrobněji Saris 2003a: 268].

MTMM metoda je velice efektivní, ale také značně náročná, neboť vyžaduje, aby byly předem do výzkumného designu zahrnuty všechny potřebné parametry. Například výzkum European Social Survey této experimentální varianty výzkumu využívá a ve svém doplňkovém dotazníku testuje přibližně 12 proměnných z hlavního dotazníku a hodnotí tak jejich validitu a reliability (více Kapitola 6.2). Metodologický test provádí třemi různými

způsoby dotazování, buď pokračuje v rozhovoru tváří v tvář, a nebo využívá formy samovyplňování dotazníku za přítomnosti tazatele nebo za jeho nepřítomnosti; tento test se provádí na třech skupinách respondentů [<http://www.europeansocialsurvey.org> (The Questionnaire)].

Odlišný přístup relativně psychologického rázu lze nalézt v aplikaci meta-analýzy, která se zabývá rozdíly mezi různými metodami sběru dat na základě jednotlivých studií, které uplatňují ve vlastních výzkumech odlišné metody sběru dat a publikují výstupy, které z výzkumů získávají. Pro všechny zkoumané studie si meta-analýza stanoví konkrétní kritéria hodnocení jejich výzkumů, například validitu použité metody, míru návratnosti, sociální desirabilitu, kterou lze z odpovědí vysledovat atp. Výsledky meta-analýzy mají pak podobu párového srovnávání způsobů sběru dat a jejich záměrem je nalézt potenciální heterogenitu mezi nimi [více o meta-analýze De Leeuw, Hox 2003: 329 – 345].

### **5.2.2 Dokumentace komparativních výzkumů**

Dostupnost dat spolu s propracovanou dokumentací výzkumů jsou součástí kvalitního vedení mezinárodního výzkumu (tzv. quality management), které činí mezinárodní výzkum transparentnější. Existují dva druhy zmíněné dokumentace: interní dokumentace projektu pro výzkumníky a dokumentace výzkumu pro sekundární analýzy dalších uživatelů dat. Prezenci dat bez řádného kontextu považuje Madden [2004: 7] za podstatu druhé největší překážky kvality komparativních analýz. Data se podle něho až příliš často potýkají s chybějícím systémem tzv. *metadat*, která poskytují informace o metodologii realizovaného výzkumu (tzv. data o datech). Metadata by měla obsahovat záznam projektu, dotazníky, výběrový vzorek, popis vyskytujících se chyb měření, oficiální statistiky atp. Tato metadata by navíc měla být rozšířena o vysvětlení nejvýraznějších rozdílů a podobností sociálního, kulturního a politického systému zemí, které jsou porovnávány. Tyto zásadní rozdíly v datech by pak měly být i v sekundárních analýzách zaznamenány v podobě viditelných varování nebo upozornění, neboť bez nich dochází k systematické, někdy až vzrůstající dezinterpretaci původních výzkumů. Dokumentace mezinárodních studií by navíc měla zahrnovat jak komplexní přehled realizace výzkumu, tak i specifickou dokumentaci každé zahrnuté země.

Nový vývoj ve standardizaci dokumentace výzkumu započal vznikem Data Documentation Initiativ (DDI), která vytvořila a prosazuje zavedení mezinárodní standardizované dokumentace dat založené na jednotné struktuře a formátu XML

(tzv. *standardní dokumentace DDI*); v dnešní době jsou tyto standardy dosazovány i do realizace národních výzkumů [Mohler, Uher 2003: 317]. DDI stanovuje kritéria obsahu, prezentace a uchovávání metadat o datovém souboru v sociálních vědách. Tyto informace pak umožňují efektivnější a přesnější využití mezinárodních dat, vylepšují proces sběru dat, šíření dat i proces datové analýzy.

Mohler a Uher [2003: 311 – 318] vymezují několik bodů, které by měla pečlivá a korektní dokumentace výzkumu dodržovat, včetně množství a charakteru údajů, které musí obsahovat :

- ◆ Vyčerpávající popis a explanaci teoretických cílů projektu.
- ◆ Dokumentaci operacionalizace teoretických jevů, včetně zmapování překladu a adaptace položek v dotazníku v různých zemích, popis originálního designu dotazníku se všemi doplňujícími materiály, designu měřících nástrojů, pilotních studií a pre-testů dotazníků.
- ◆ Popis výběrové procedury a její realizace (definice populace a její velikost, detaily monitorování terénního sběru dat, časový harmonogram, instrukce pro tazatele a charakteristika tazatelské sítě, popř. zmapování agentur zajišťujících sběr dat a financování výzkumu).
- ◆ Popis finální podoby datového souboru, tzn. výstupy výzkumu – míra návratnosti, pokusy o kontaktování a popisy nerespondentů, chybějící hodnoty v datovém souboru, přehled o kódovaných proměnných a filtrech, odhady výběrové chyby a její zdroje, popis vážení, editování dat. V rámci mezinárodního výzkumu se při popisu datového souboru vytváří tzv. *meta-codebook*, který spojuje datový soubor s původním dotazníkem a kulturně specifickými otázkami na mezinárodní úrovni (obsahuje metadata) [Mohler, Uher 2003: 318]. Editování dat je v mezinárodních projektech prováděno na národní úrovni a nebo se provádí centrálně na jednom místě, kde analytici hledají chyby v datovém souboru, vyřazují diskrepance mezi daty, zajišťují logickou konzistenci dat a data harmonizují. Proces harmonizace v podstatě reaguje na rozdílnosti ekvivalentních pojmů a transformuje položky v dotazníku do odvozené komparativní proměnné, kterou již lze využít v analýzách [Mohler, Uher 2003: 326].
- ◆ Detailní popis sociokulturního a politického kontextu mezinárodní studie a rozdílnosti terénů sběru dat (popis zahrnuje podle potřeb konkrétního šetření a analýzy základní populační statistiky i specifické proměnné jako je např. míra rozvodovosti, regulace potratů, azylová politika atp.).



### 5.3 Psychometrické přístupy v rámci dosahování kvality šetření

Psychometrické přístupy v rámci dosahování kvality mezinárodních komparativních výzkumů zasahují do celé konstrukce nástroje měření. Právě aplikací různých psychometrických technik se lze vyhnout nebezpečí radikální operacionalizace zkoumaných jevů, která zabraňuje dosažení funkční ekvivalence pojmů v dotazníku [Przeworski, Teune 1970: Part II.; Nowak 1989]. Tyto přístupy představují ucelený zdroj faktorů, jež mají vliv na použití dotazníků pro mezinárodní výstupy, a jsou součástí většiny (nejen) mnohojazyčných studií zabývajících se komparativními šetřeními.

#### 5.3.1 Způsob vytváření dotazníků a jejich překlad

Metody vytváření dotazníků pro použití v různých kulturách svým charakterem typicky spadají do psychometrických přístupů. Sociální vědci se v této oblasti zabývají převážně možnostmi vytváření designu mezikulturních dotazníků, srovnáváním současných typů dotazníků a vývojem nových dotazníků, nebo problematikou převzetí (*adopting*) a adaptace (*adapting*) již existujících položek měření [Harkness, Van de Vijver, Johnson 2003: 25].<sup>10</sup>

S ohledem na strukturální problémy s překladem významů pojmů do optimální verze, kterou lze použít ve všech zemích mezinárodního projektu, se klade největší důraz na relevantní, precizní a *ekvivalentní překlad* dotazníků. Je to nejčastěji zmiňovaný problém v literatuře věnující se empirickým komparativním výzkumům. Hraje důležitou roli jako jedna z pěti základních procedur ve vytváření finální verze dotazníku, tzv. procedura **TRAPD**: **T**ranslation (překlad), **R**eword (revize), **A**djudication (posouzení), **P**retesting (předběžné testování), **D**ocumentation (dokumentace) [Harkness 2003: 38].

Lingvistické a kulturní faktory spojené s přítomností odlišných zemí mezinárodních šetřeních vždy citelně ovlivňují, z tohoto důvodu nemohou být ignorovány, protože výzkumníci tak riskují ztrátu kontroly nad konstrukcí a analýzou klíčových proměnných a potažmo dosažením celkové kvality výsledků projektu. Jowell [1998] uvádí, že každá druhá země musí dotazník přeložit do vlastních funkčně ekvivalentních slov a frází (většinou z oficiální britské angličtiny) a tento proces s sebou přináší nemalé problémy.

Kupříkladu většina anglicky hovořících zemí (např. Austrálie, Nový Zéland, Kanada, Irsko, USA) se v rámci mezinárodních výzkumů vůbec neobtěžuje předložené dotazníky

---

<sup>10</sup> Pro vývoj nových dotazníků se například používají 3 různé přístupy: sekvenční (Eurobarometer), paralelní (ISSP) a simultánní (WHO) [více Harkness, Van de Vijver, Johnson 2003: 25].

z britské angličtiny překládat, čímž činí množství vážných chyb. Nehledě na to, ani ostatní participující země by se neměly smířovat s přímým, jednosměrným nebo doslovným překladem a využívat služeb pouhého jednoho překladatele (tzv. one-to-one translation). Nejznámější překladatelský přístup používaný v mezinárodních výzkumech je tzv. *technika zpětného překladu* (back translation model), která velmi pečlivě kontroluje validitu překladu [Harkness 2003: 41]. Základní logika zpětného překladu spočívá v přeložení originálního (většinou anglického) dotazníku do rodného jazyka dané země. Poté požaduje zpětný překlad přeložené verze dotazníku do původního anglického jazyka, porovnat jej s originálem, a opět pokračovat v procesu překládání z této druhotně vytvořené verze anglického dotazníku. V procesu překládání se má takto postupovat až k dosažení nejlepšího, nejbližšího možného výsledku. Navíc každý tento krok překladu má za povinnost provést jiný překladatel.<sup>11</sup>

Proces zpětného překladu by měl podchytit zásadní překladové odchylky, ale Johnson [1998] varuje před falešným pocitem bezpečí. I v současné době mezinárodní komparativní programy (např. ISSP) velice vzácně zaměstnávají odborné překladatele, převážně z finančních důvodů, oproti tomu ale nesmírně často spoléhají ve všech členských zemích na odbornou znalost sociálních vědců účastnících se daných výzkumů. I když se do výzkumného procesu zapojí bilingvní překladatelé s hlubokou a aktuální znalostí obou kultur, vždy je na místě zvažovat jejich potenciální nedostatek znalostí metodologických požadavků mezikulturního překladu [Johnson 1998: 18]. Dále se musí vést v patrnost, že jazyk je rovněž citlivý k sociální vrstvě překladatele i adresáta (resp. respondenta), kterému bude překlad doručen v podobě dotazníku [Disman 2005].

V tomto kontextu se pak vynořující fascinující odlišnosti různých zemí stávají velice lehce iluzorní, neboť pocházejí z lidské chyby nejistého překladu, kde stačí, aby se přeložené slovo nebo fráze chovaly jako jemně odlišné položky od těch původních. Tento problém se nejčastěji zvažuje v možnostech mezinárodního překladu přibližně stejných odpovědí v dotaznících. Jowell [1998] na toto téma pokládá provokativní otázku: „*Najdeme dobře sedící a jasně srozumitelný funkční ekvivalent slovních spojení 'spíše souhlasím, trochu souhlasím, poněkud nesouhlasím, spíše nesouhlasím' v polštině, japonštině, maďarštině a bengálštině?*“

Dnes již existují různé variace a vylepšení tohoto modelu zpětného překladu

---

<sup>11</sup> Proces překladu v projektu European Social Survey postupuje podle specifického protokolu, je pečlivě zaznamenáván každý krok a jednotlivé rozhodovací úseky pro budoucí citace, odkazy a diskuze. Finančně zajištění by měli být standardně tři zkušení překladatelé pro překlad, kontrolu a úpravu překladu a provedení dvou nezávislých překladů. Pre-testování zvolené formy dotazníku se provádí na zhruba 50 lidech vybraných klasickým kvótním výběrem.

(např. decentering back-translation), využívá se několika alternativních verzí překladu dotazníku, náhodného použití různých verzí dotazníku ve vybraných vzorcích respondentů a zemí, včleňují se do procesu překladu poznatky kognitivních přístupů [Johnson 1998: 20]. Rozsáhlé mezinárodní programy navíc také požadují poskytnutí doprovodných poznámek k pojmovému systému dotazníku a podrobné popisy významů potenciálních mnohoznačných otázek.

### 5.3.2 Vývoj komparabilních otázek a odpovědí

Konstrukce multikulturních dotazníků, které jsou funkčně ekvivalentní ve všech cílových populacích, je základní problém kvality mezikulturního výzkumného měření. Otázky musí být samostatně validní a navíc musí splňovat také srovnatelnou mezinárodní validitu. Tato složitá situace se netýká jen oblasti dotazování, ale i zaznamenávání odpovědí v dotaznících. Jacobson, Kumata a Gullahorn [1960] poznamenávají: „*Obtížnosti vznikají již v základních teoretických otázkách týkající se psychologických procesů, které zasahují mezi měřené položky a odpovědi v monokulturních národních studiích, mezinárodní výzkumy tuto obtížnou situaci ještě umocňují a přidávají na vrch další nové problémy.*“ Velké rozdíly v jazycích, kultuře i sociálních strukturách jednotlivých zemí, které činí mezinárodní výzkumy tak analyticky hodnotné, jsou zároveň i zásadní překážkou v dosahování empirické ekvivalence měřených položek [Smith 2003: 73].

Metodologickým rámcem vytváření otázek v dotazníku je oblast kognitivních přístupů, která se problematikou psychologických stránek získávání informací od respondentů ve standardizovaných dotazníkových šetření zabývá [Vinopal 2008: 11]. V centru pozornosti kognitivních přístupů stojí problematika interpretace otázky na individuální a interpersonální úrovni. Mapují, jakým způsobem respondent na výzkumné otázky reaguje, jak chápe a zpracovává položky v dotazníku. Také identifikují reálné či potenciální zdroje zkreslení odpovědí na úrovni interakce a komunikace respondenta s tazatelem.

Na mezinárodní úrovni komparativních výzkumů se hlavní zájem kognitivních přístupů obrací na přítomné kulturní kontexty, které ovlivňují interpretaci položených otázek a formování odpovědí [Braun 2003a: 57]. Základní otázka v tomto smyslu zní: Jak respondenti *z různých kultur* rozumí dané otázce? A zatímco okolnosti interpretace otázky jsou řešeny zpracováváním technických standardů srovnatelnosti a ekvivalence položek v dotazníku (viz. Kapitola 4), otázka úpravy odpovědí a odpovědních škál se stává ústředním tématem mnoha odborných studií tohoto zaměření.

Zásadní problém tkví v překladové ekvivalenci škál odpovědí, v efektu kontextu nabídky odpovědí, v preferenci jedné odpovědní kategorie před jinými a jejich různé interpretaci a ve formálním způsobu dotazování. Například neverbální škály odpovědí, které používají vizuální pomůcky, nemusí být ekvivalentní mezi různými kulturami. Jiné, ordinální škály využívají stejného řazení kategorií odpovědí na škále bez ohledu na konkrétní kulturu, ve které jsou používány. Navíc mohou mít podobu reálně věcných škál s odpověďmi „denně, týdně, měsíčně“, nebo psychologických škál v podobě „často, čas od času, zřídka“. Kategorie v psychologických termínech nemají ale stabilní význam vzhledem k daným kulturám, samotné posouzení odpovědi v sobě zahrnuje srovnávání sama sebe s ostatními členy skupiny nebo s normami dané skupiny, což pro komparaci mezi zeměmi není vhodné [Przeworski, Teune 1970: 98]. Týdenní čtení novin například připadá univerzitnímu profesoru v Americe řídké, asijský zemědělec může ale považovat týdenní čtení novin za časté. I numerické ordinální škály, které používají místo slovních kategorií jako odpovědi čísla, nejsou zcela ideální. Toto měření opět není standardní pro všechny země. Pokud má respondent vybrat číslo na desetibodové škále, význam, který přiřazuje například číslu šest, není stejný ve všech zemích a řazení se tak opět stává subjektivní a nesrovnatelné.

Jako velice robustní a reliabilní se naopak jeví tzv. kalibrované škály, které se vyhýbají absolutně definované hodnotě kategorie odpovědi a ani nedefinují vzdálenosti mezi nimi. Respondenti sami řadí sílu termínů odpovědí a definují si body předložené škály, jež se později v analýze převedou do jejich číselných ekvivalentů (např. alfabetská škála, sémantický diferenciál atp.) [Smith 2003: 77].

Saris [2003b: 275] v otázce získávání a zaznamenávání odpovědí argumentuje tím, že proces, kterým respondenti vybírají odpovědi mezi nabízenými kategoriemi, se mezi kulturami vždy liší, a že odpovědi v podstatě nejsou dobrým indikátorem respondentova úsudku. Nejpodstatnější je podle něj v tomto okamžiku určit pravděpodobnosti odpovědí v různých zemích (response probabilities).

Celá oblast problematiky efektu odpovědi (response effect) je v mezinárodním vnímání vztažena také k otázce *sociální desirability*. Sociální desirabilitu by bylo možno definovat jako určitou potřebu sebe prezentace, která vede respondenty k zobrazování sama sebe v lepším světle, a souvisí s nadhodnocováním všeobecného mínění a sociálně akceptovatelného chování. V podstatě je otázka sociálně desirabilní tehdy, pokud odklání respondenty od variant odpovědí, které jsou považovány za společensky nežádoucí [Vinopal 2008: 47].

Tento jev je diskutován ze dvou úhlů pohledu, za prvé z hlediska fungování sociální desirability jako stylu odpovědi respondenta (response style), z druhé perspektivy se sociální desirabilita vnímá jako osobnostní psychosociální charakteristika [Johnson, Van de Vijver 2003: 195]. I když se sociální desirabilita ukazuje jako univerzální koncept, který se objevuje ve všech sociálních skupinách, obsahuje kulturně specifické faktory a má různý vliv na respondenty v různých kulturách a sociálních skupinách. Je například dokázáno, že skupiny s menší (relativní) sociální mocí či pozicí ve většinové společnosti mají větší tendenci k sociálně desirabilnímu chování [Ross, Mirowski 1984]. Některá témata jsou ve všech sociálních skupinách systematicky podhodnocována (např. kriminální jednání, užívání drog, konzumace alkoholu, některé druhy příjmů atp.), naopak některá jsou systematicky nadhodnocována (volební účast, šetření energií atd.). Tyto případy uvádějí Tourangeau, Rips a Rasinski (2002) za empiricky prokázané [podrobněji Vinopal 2008: 47].

Vliv sociální desirability lze pocítit v procesu odpovědi na otázku i při komunikaci respondenta s tazatelem v závislosti na citlivosti konkrétního tématu, která je ještě navíc různě kulturně posuzována. Efekt sociální desirability může být tedy ovlivněn i zvolenou metodou sběru dat, proto je mu v mezinárodních výzkumech věnována tak velká pozornost.

#### **5.4 Prostředky zmírňující problémy srovnávání (rady a doporučení)**

Mnoho autorů a metodologů na všechny tyto komplikace kvality mezinárodního srovnávání reagují různým způsobem a následně se pokouší formulovat řadu doporučení pro zmírnění dopadu konkrétních problémů v komparaci dat (viz. Hantrais, Mangen 1996; Jowell 1998; Madden 2004).

Jowell [1998] načrtnul několik obecných pravidel zmírňujících některé problémy mezinárodního srovnávání. Sociální vědci by se podle něj měli v první řadě zavázat k tomu, aby neinterpretovali výzkumná data vztahující se k zemi, o níž nic nevědí a nebo jen velmi omezeně, a to jak z hlediska tradice používaných metod, tak i sociální struktury studovaných zemí a jejich historického a kulturního zázemí. Splnění tohoto požadavku zároveň vede k zajištění mezinárodní spolupráce v interpretaci potenciálních kulturních rozdílů.

Analytici mezinárodních dat by měli také odolat pokušení srovnávat příliš mnoho zemí najednou, neboť tím narůstá heterogenita vzorku. Měli by se také pokusit srovnávat země, které jsou si blízké a něčím podobné. V tomto kontextu jde o klasický požadavek objektivity, musí být zcela zřejmé, z jakého důvodu je komparace prováděna. Jowell [1998]

dále upozorňuje na fakt, že čím více zemí bude zahrnuto do studie, tím méně jich bude ve stejném měřítku participovat na rozvoji výzkumné práce (např. pilotních studií, pre-testování výzkumných nástrojů). Navíc by měli být výzkumníci schopni ze svých srovnávacích analýz vypustit země s nízkou návratností, nepřizpůsobivou prací v terénu, země, které používaly odlišné způsoby dotazování nebo rozdílné metody výběru.

Analýza by měla mít přirozeně stanovené a jasně vymezené hranice. Pouze některé sociální jevy a pouze některé aspekty těchto jevů mohou být úspěšně měřeny mezinárodně. A i tyto jevy vhodné ke srovnávání vyžadují velké kompromisy pro svá vymezení a operacionalizaci, která zajistí jejich konzistentní měření v různorodých zemích a kulturách. Navíc by se měli analytici zdržet svého prvotního přesvědčení o nalezení odlišností mezi národy a být skeptičtí. Až příliš mnoho na pohled validních odlišností je výsledkem špatného překladu či například jemně odlišné obrazové přílohy dotazníku, nebo různých variant rozsahu výběru, a jsou tedy jen nekonzistentními zjištěními bez větší váhy. Může být značně nerozumné vytvářet radikální a poutavé závěry na základě neočekávané národní odchylky v odpovědi na jedinou otázku [Hantrais, Mangen 1996].

Aby byla zajištěna kvalita získaných dat, musí být stanoveny rovnocenné a přísné technické standardy pro komparativní výzkumné metody (rozsah vzorku, výběrové metody, výpočet návratnosti dotazníků, práce v terénu), bezpodmínečně platné pro každou zemi, která se mezinárodního projektu účastní. Až příliš často analytici také srovnávají soubory národních dat v jakémisi vzduchoprázdnu, pozornost by se měla věnovat sestavení celkové úrovně proměnných, které mohou být různým sociálním kontextem výzkumu ovlivněny (tzv. kulturní nebo sociální interference). Například v rámci tématu náboženství musí být do datového souboru zahrnuty celkové informace o struktuře a zvycích hlavního náboženství, církví a sekt uvnitř každé země [Jowell 1998].

V otázce kvalitní analýzy mezinárodních dat Denise Lievesley [2001: 378] navrhuje badatele k vyhnutí se vágním srovnávacím tabulkám (tzv. league tables). Tyto kategorizační tabulky, které jen řadí velké množství zemí podle účelově vybraných proměnných, směřují k ploché interpretaci bez kontextu, který je zapotřebí. Navíc, pokud srovnávací tabulky využívají nestandardizovaná data z různých výzkumů odlišných zemí, je tato komparace značně nesolidní. Někdy může být tento typ výstupů dat zneužit k manipulaci informací, proto je jejich využívání v současné době věnována speciální pozornost. Pro konstrukci interpretačně zajímavých analýz se také doporučuje využít data, která nějakým způsobem vystihují trendy. Trendy by měly být totiž daleko méně ovlivněny metodologickými a kontextovými rozdíly,

za předpokladu, že změny zachycené v datech jsou stabilní v čase a všechny vnější vlivy zahrnuté do analýzy. Například není rozumné porovnávat úroveň finanční podpory ministerstev mezi zeměmi pro jakýkoli konkrétní rok, ale trendová data už mohou poskytnout daleko reliabilnější indikátor pro vzrůstající a klesající státní podporu v pozorovaných zemích [Madden 2004: 10].

Ve srovnávacích analýzách by se výzkumníci ani neměli vyhýbat použití různých poměrů proměnných a jiným statistickým ukazatelům. Poměry a jiná statistická měření mohou totiž předejít mnoha problémům vznikajícím při srovnávání. Například poměr umělcova příjmu vůči průměrnému příjmu obyvatele může být daleko užitečnější komparativní statistika než pouhý zprůměrovaný příjem umělců ve všech zemích. Poměr umožňuje v důsledku dosáhnout umělcova relativního příjmu napříč zeměmi a vyhýbá se potřebě vybrat vhodný směnný kurz [Madden 2004: 10].

## 6 Komparativní kvantitativní výzkum European Social Survey

Situace v daném výzkumném prostředí významně určuje, jaké jsou priority výzkumníků při přípravě a realizaci designu kvalitního šetření. O odlišnostech podmínek v jednotlivých zemích svědčí například praxe mezinárodních šetření, která odhaluje, jak velmi omezené jsou možnosti aplikace srovnatelných metodik. Tyto velké rozdíly výzkumných prostředí různých kultur jsou patrné z analýz metodik a kvality šetření mezinárodních programů a odrážejí se i v problémech se stanovením metodických pokynů pro realizaci těchto programů [Krejčí 2008: 43].

V zájmu detailnějšího průzkumu současné podoby mezinárodních programů (tzv. *Cross-National Survey Programs* - ESS, ISSP, WVS, EVS, EB atp.), dodržování požadavků kvalitní metodologické přípravy a vyrovnávání se s teoreticko-empirickými problémy mezinárodních výzkumů v praxi, je vhodné věnovat hlubší pozornost pouze jednomu zvolenému zástupci mezinárodních projektů. Tímto ukázkovým příkladem bude nadále jeden z nejnovějších současných projektů sociálních šetření organizovaných v Evropě v akademické sféře: *The European Social Survey* (ESS).

### 6.1 Představení projektu The European Social Survey (ESS)

European Social Survey je významný kvantitativní kontinuální výzkumný projekt v oblasti sociálních věd, na jehož počátku bylo poznání, že ve většině evropských zemí chybí relevantní a kvalitní data, která by umožňovala zkoumat postoje a hodnoty v evropských společnostech v časové a mezinárodní komparativní perspektivě. Z důvodu snahy o získání standardizovaných a komparabilních dat si dal projekt ESS za cíl obsáhle se zabývat populacemi všech demokratických zemí západní, střední a východní Evropy [Plecitá-Vlachová, Kalvas 2005].

Už od roku 2001 ESS každé dva roky mapuje dlouhodobé postoje a změny v chování v kontextu evropského sociálního, politického, kulturního a morálního klimatu. ESS se věnuje výzkumu politických témat a problémů (kriminalita a nerovnosti atp.), postojům k demokracii (spokojenost s demokracií, legitimita demokracie, důvěra v instituce, politická orientace atd.), politické participaci a volebnímu chování, sociální a politické orientaci (individualizace, sociální vztahy, viktimizace, postmaterialismus atp.), využívání masmédií a komunikaci (čtení novin, poslech rádia, sledování TV, užívání internetu atd.), sociálně politické identitě (národní identita, migrace, etnocentrismus, xenofobie, patriotismus, postoje k EU atd.), prosperitě,



spokojenosti se životem, hodnotovým žebříčkům, náboženství a jiným tématům. V šetřeních jsou v jednotlivých zemích sbírány socio-demografické proměnné (věk, pohlaví, rasa/etnikum, občanství atp.) a socio-strukturní charakteristiky (sociální pozice, vzdělání, pracovní život, sociální exkluze atp.) dotázaných a jejich domácností.

Sebraná data mají za úkol odhalit neobvyklé kontrasty napříč 30 zeměmi, většiny zemí Evropské unie, řady dalších evropských zemí a některých mimoevropských zemí (Turecko, Izrael, Rusko). Kromě toho ESS poskytuje prostředky, díky kterým mohou země posuzovat a hodnotit samy sebe, také částečně podle toho, jak se jejich obyvatelé cítí nebo jak se vyslovují ke své aktuální existenční situaci.



Zdroj: převzato z  
<http://ess.nsd.uib.no>

První vlna výzkumu ESS proběhla v letech 2002 – 2003, účastnilo se jí 22 zemí. Druhá vlna 2004 – 2005 s účastí 26 zemí. Třetí vlna 2006 – 2007, 25 zemí. Čtvrté vlny 2008 – 2009 se účastnilo 31 zemí (Pátá vlna se chystá na rok 2010 – 2011): Belgie, Bulharsko, Chorvatsko, Česká republika, Dánsko, Estonsko, Finsko, Francie, Holandsko, Irsko, Island, Itálie, Izrael, Kypr, Litva, Lotyšsko, Lucembursko, Maďarsko, Německo, Norsko, Polsko, Portugalsko, Rakousko, Rumunsko, Rusko, Řecko, Slovensko, Slovinsko, Španělsko, Švédsko, Švýcarsko, Turecko, Ukrajina, Velká Británie

I když tento evropský projekt sdílí jistou podobnost s jinými velkými dlouhodobými mezinárodními komparativními programy (viz. EVS, ISSP, EB atp.), v žádném případě je nenahrazuje [Plecitá-Vlachová, Kalvas 2005]. Tento exaktní komparativní výzkum je využíván k informování a doplňování akademických a politických debat. Poskytuje badatelům, zákonodárcům evropských vlád, agenturám a ostatním zainteresovaným skupinám možnost opětovně měřit a interpretovat změny v lidských hodnotách a postojích. Je to první sociálněvědní projekt, který získal prestižní cenu „Descartes Prize“ (2005) za excelentní vědecký výzkum prováděný ve spolupráci vědců. Oceněn byl jejich postup na poli komparativních sociálních měření systematickým využíváním přísných metodologických a akademických standardů. Akceptováním striktních přístupů ohledně pravděpodobnostního vzorku, testování otázek, zaznamenávání kontextových událostí, precizního překladu a vylepšení míry návratnosti se ESS stal směrodatným a oficiálním zdrojem informací, týkajících se změn sociálních hodnot v transformující se Evropě.

Klíčovým cílem ESS není výzkum konkrétního tématu, ale postupné budování databáze časově a mezinárodně srovnatelných indikátorů, snadno a rychle dostupných všem

výzkumníkům pro různé výzkumné účely (dlouhodobé sledování změn, řešení náročných analytických úloh atp.). Podmínky přístupu do databáze jsou stejné pro všechny zájemce včetně autorů šetření. ESS tak sám vlastně není výzkumem, ale vytváří jednu z nejvýznamnějších infrastruktur pro evropský sociální výzkum a jeho data nacházejí uplatnění v nejrůznějších sekundárních projektech. [Krejčí 2008: 92]

Ambiciózní projekt ESS byl podnícen Stálým výborem pro sociální vědy v Evropské nadaci pro vědu (European Science Foundation ESF) a spuštěn díky finanční podpoře Evropské komise (European Commission) z rámcových programů a Evropské nadace pro vědu, která hradí náklady spjaté se základní existencí programu (vývoj výzkumných instrumentů, schůzky orgánů ESS, archivování a distribuce dat, atp.). Zúčastněné země hradí realizaci terénních výzkumných šetření realizovaných v rámci ESS na jejich území (nejčastěji skrze národní akademické organizace) z národních zdrojů.<sup>12</sup>

## 6.2 Metodologické standardy projektu ESS

Projekt ESS je navržen a prováděn tak, aby splňoval nejvyšší metodologické standardy uplatňované v této oblasti kvantitativních šetření. Mezi cíle ESS patří vysoká vědecká úroveň metodiky a časová a mezinárodní srovnatelnost výstupů. ESS proto usiluje o standardizaci a redukci rozdílů designu šetření v jednotlivých zemích a požaduje dodržování poměrně podrobně vymezených metodických postupů a parametrů šetření, přičemž navržené metodiky jednotlivých šetření musí odpovídat podmínkám v dané zemi [Krejčí 2008: 92]. Konkrétní metodické návrhy jednotlivých zemí předem schvaluje mezinárodní Centrální koordinační tým (viz. Příloha 1).

Mezi nejrozšířenější používaná kritéria pro stanovení kvality šetření patří reliabilita, validita, rozsah chyb návratnosti, relativní systematická chyba, efekt odpovědi, neporozumění otázce a problémy v interakci mezi tazatelem a respondentem [ESS 2002a]. V posledních letech se dělají v mezinárodním projektu ESS významné pokroky v dosahování větší ekvivalence výzkumu zvláště po stránce výběru vzorku, měření míry návratnosti dotazníků, způsobu práce v terénu, zajišťování dat a jiných oblastí [Fitzgerald 2009]. Avšak klíčová oblast, jež zaostává za úspěchem rovnocennosti výzkumů v technické oblasti realizace, je základ pro rozvoj efektivní dotazníkové práce a mezinárodních ekvivalentních dotazníků [ESS 2002a].

---

<sup>12</sup> Podrobnější informace o tomto programu lze nalézt na: <http://www.europeansocialsurvey.org>, informace poskytuje R. Jowell (ESS coordinator): *About the European Social Survey*

#### A) Podoba dotazníkového šetření

European Social Survey upřednostňuje realizaci osobních rozhovorů tazatelů s respondenty (face-to-face) přibližně délky jedné hodiny ve všech zúčastněných zemích. ESS je ale také zároveň pro flexibilnější variantu sběru dat tzv. „*mix-mode design*“ (viz. Kapitola 5.2), proto umožňuje participujícím zemím výběr buď z papírové formy dotazníku, telefonického rozhovoru nebo elektronického výzkumu podle jejich vlastních výzkumných potřeb a zvyklostí. První kontakt tazatele s respondentem je tedy vždy osobní (face-to-face), s výjimkou země s oporou výběru jmen konkrétních lidí s telefonními čísly, ti se primárně kontaktují telefonicky. Dotazník sestává z hlavního dotazníku a doplňkového dotazníku (*supplementary questionnaire*), který v rámci rozhovoru s respondentem následuje až po hlavním dotazníku.

Každý hlavní dotazník má dvě části, základní modul otázek (core questionnaire) a zpravidla dva až tři rotační moduly (rotating questionnaire). Hlavní položky dotazníku pokrývají socio-demografické údaje a dvě třetiny zkoumaných témat, obsah těchto položek zůstává prakticky stejný v každém opakovaném výzkumu. Základní modul zahrnuje závislé i nezávislé proměnné navržené tak, aby měřily změny v čase, které jsou považovány za klíčové komponenty evropské sociální struktury. Stabilní hlavní položky jsou vytvářeny ve spolupráci s expertními skupinami z rozdílných odvětvích sociálního, politického a ekonomického života a pokrývají již zmíněná hlavní témata.

Střídající se položky v dotazníku z rotačních modulů nejsou konstantní a vycházejí z požadavků mezinárodního týmu sociálních vědců z celé Evropy. Procházejí variantou výběrového řízení, kde nejúspěšnější návrh je ten (v rámci relevance a časového rozvržení), který má šanci vytvořit největší příspěvek současné sociální vědě a politice v evropském kontextu. Například v první vlně ESS (2002) byl poskytnut prostor rotačním modulům „Občanství, občanská angažovanost a demokracie“ a „Imigrace“. V současné době se překládají už jen střídající se položky, hlavní dotazník je už stabilně přeložený.<sup>13</sup> Každá země si navíc může do dotazníku ESS přidat jakékoli vlastní otázky pro specifické národní využití, avšak tyto otázky musí být kladeny až po všech otázkách hlavního a doplňkového dotazníku a jejich počet i rozsah musí být předem konzultován s Centrálním koordinačním týmem, obzvláště kvůli záruce nenarušení návratnosti dotazníků.

Doplňkový dotazník má za úkol za prvé mapovat otázky základních lidských hodnot. Nástroj, který byl pro tento účel vyvinut, je měřicí škála 21 položek týkajících se proměny

<sup>13</sup> Aby zajistilo ESS náležité pokrytí populace, překládá dotazníky do všech hlavních jazyků participujících zemí, což znamená, že jimi hovoří vždy více jak 5 % obyvatel dané země.

lidských hodnot (Human Value Scale), kterou navrhl izraelský psycholog profesor Schalom Schwartz. Za druhé má doplňkový dotazník testovat proměnné z hlavního dotazníku opakovaným dotazováním jemně odlišných forem stejných otázek. Tento druh testu hodnotí reliabilitu a validitu vybraných stimulů z hlavního dotazníku a zjišťuje chybu měření. Využívá MTMM designu (viz. Kapitola 5.2), tudíž provádí metodologický test třemi různými způsoby dotazování. Doplňkový dotazník buď pokračuje v osobním dotazování a nebo je vypracováván formou samovyplňování za přítomnosti tazatele nebo samovyplňováním za nepřítomnosti tazatele; v tomto případě následný sběr vyplněných dotazníků probíhá poštou či osobně. Tento experimentální design se provádí na třech skupinách respondentů po 12 otázkách.

#### *B) Výběr vzorku*

Kvantitativní výzkum je založen na premise, že „*odpovědi několika respondentů jsou reprezentativní pro všechny*“. Tento předpoklad je pravdivý pouze v případě, pokud je vzorek respondentů reprezentativní v demografických charakteristikách stejně jako ve vzorcích chování nebo postojích pro celou cílovou populaci. Tento problém má daleko větší váhu v komparativních výzkumech, kde mohou rozdílné odpovědi mezi zeměmi ohrozit celkovou validitu mezinárodního srovnávání [Couper, De Leeuw 2003].

Z tohoto důvodu ESS zavádí výjimečně vysoké standardy pro výběr vzorku populace. Cílem výběrového výzkumného plánu je plné pokrytí rezidenční populace za použití pouze náhodných pravděpodobnostních výběrových metod ve všech krocích výběru. Není povolena žádná náhrada respondenta v žádném kroku výběru. Minimální velikost požadovaného výběrového vzorku (effective sample size) je většinou 800 respondentů pro země pod 2 miliony obyvatel, 1500 respondentů pro všechny ostatní země s vyšším počtem obyvatel, aktuální počet respondentů ve výběru však může být vyšší.

Výzkum musí být reprezentativní pro všechny osoby nad 15 let obývající vlastní domácnosti a to v každé zemi, bez ohledu na národnost, občanství a jazyk. V předvýzkumu musí každá země ve výběrovém panelu přesně zjistit vhodnou velikost vzorku do výběru (podle minimální požadované velikosti a designového efektu), v úvahu se musí vzít i realistický odhad vlivu klastrování, vhodnost měř a vážení a míra návratnosti response rate. Každé potenciální nepokrytí určité skupiny (z důvodů potíží s jazykem, nedostatků v rámci výběru atp.) musí být konfrontováno s předešlým výběrovým panelem kvůli konečnému rozhodnutí ohledně výběrové metody, která dané problémy vyřeší.

ESS požaduje co nejkvalitnější pravděpodobnostní výběr. Za tímto účelem byl

v jednotlivých zemích proveden průzkum dostupných opor (seznam adres a domácností, populační statistiky, seznam obyvatel atp.) a podmínek pro realizaci pravděpodobnostních šetření se snahou vytipovat nejvhodnější metodu [Krejčí 2008: 61]. Jednotlivé země lze podle použitých metodik rozdělit do čtyř skupin. První tvoří poměrně velká skupina zemí, kde jsou k dispozici obvykle velmi kvalitní údaje z populačních registrů (např. Dánsko, Švédsko, Finsko, Norsko), ve druhé skupině jsou země s dostupnými databázemi domácností (např. Španělsko, databáze SIPO v České Republice), třetí skupinou tvoří země s databázemi adres (Nizozemsko, UK) a čtvrtou skupinou státy, kde není k dispozici žádná opora (Francie, Portugalsko, Řecko) [ESS 2004].

### *C) Design výběru a jeho efekt*

Výběrové principy ESS dodržují Kishovu myšlenku (1994) o flexibilitě designu výběru za předpokladu pravděpodobnostních výběrových metod. Volba specifického designu výběru závisí na vhodném rámci, zkušenostech a také ceně v jednotlivých zemích. Srovnatelnost a kompletní pokrytí osob z cílové populace jsou hlavní cíle výběrové strategie [ESS 2007]. Místo prostého náhodného výběru tak většina zemí v projektu používá komplikované metody víceúrovňových výběrů a jiné složitější výběrové plány např. skupinové výběry. Výběr proto není zcela nezávislý a jednotky ve skupinách jsou si zpravidla navzájem podobnější, rozptýl se ve srovnání se situací prostého náhodného výběru liší a efekt velikosti vzorku se snižuje [Krejčí 2008: 63].

Tyto komplikované designy šetření mají často problémy s dosahováním přijatelné velikosti výběrové chyby. Velké rozdíly v pravděpodobnosti výběru jednotlivých jednotek ve všech krocích výběru vedou k velkému *designovému efektu*, který je způsobený konkrétním výběrovým plánem [Ganninger 2006; Krejčí 2008: 63].

Hodnota designového efektu má udávat, kolikrát je rozptýl větší, než kdyby se jednalo o prostý náhodný výběr. Tento prostý náhodný výběr má takovou velikost vzorku (tzv. effective sample size), která vyprodukuje stejně přesné měření (resp. standardní chybu) jako design aktuálně použitého výběru. Typicky je velikost effective sample size menší číslo než aktuální počet sebraných respondentů. Pokud chceme dosáhnout u aktuálně realizovaného výběru stejné statistické průkaznosti jakou má prostý náhodný výběr, musíme buď zvětšit efektivní velikost vzorku tolikrát, kolik udává hodnota designového efektu, nebo naopak uvažovat větší rozsah konfidenčních intervalů [Krejčí 2008: 63].

Hodnota celkového předpokládaného designového efektu spočtená metodikou ESS se následně odráží na teoretické práci s daty, se kterými by se mělo hypoteticky pracovat s n-krát

větší výběrovou chybou než v případě prostého náhodného výběru (např. 2x), velikost výběru je tomu pak přizpůsobena a je provedeno n-krát více rozhovorů (např. místo 1500 je to 3000 rozhovorů) [Krejčí 2008: 63]. Designová váha (deff) musí být vypočítána předem, na základě dvou datových souborů z předešlé vlny výzkumu, pro aktuální výběrový design kvůli stanovení počtu respondentů. Pokud designový efekt v jednotlivých zemích v průběhu let klesá, země dokázaly zkonstruovat výzkum lépe. Například Norsko přešlo ze stratifikovaného výzkumu na prostý náhodný výběr a dokázalo tak podstatně vylepšit svou designovou váhu [Krejčí 2008: 63].

#### *D) Návratnost*

Za cíl si ESS stanovuje minimální návratnost dotazníků na 70%, pokud jsou doplňkové dotazníky distribuovány odděleně od hlavních dotazníků, jejich minimální návratnost je 90%. Dalším přísným pravidlem je udržet minimální výskyt nekontaktovatelných respondentů (non-contacts) na 3 %. Zpětná kvalitativní kontrola (osobně, poštou či telefonem) musí být provedena a zdokumentována na 5% respondentech, 10% respondentech odmítajících odpovědět a 10% nekontaktovatelných respondentech. Definice metodik pro všechny vlny ESS proto zahrnovaly následující opatření pro zvýšení návratnosti dotazníků:

- ◆ sběr dat alespoň po dobu jednoho měsíce a jeho monitoring
- ◆ řádné školení všech tazatelů a vylepšování získávání odpovědí
- ◆ omezené pracovní zatížení tazatele (max. 24 jednotek výběru, tazatel nesmí provést více než dva tyto tazatelské úkoly)
- ◆ zpětná kontrola rozhovorů
- ◆ nejméně 4 výzvy (telefonní hovory, krátké osobní návštěvy) pro každou výběrovou jednotku předtím, než je označena za non-response, výzvy mají být provedeny v různé dny a v různé časy, aby skutečně lokalizovaly potenciální respondenty, jedna výzva ráno v pracovní den a jedna o víkendu ve dvou různých týdnech
- ◆ zákaz náhrady respondentů ve všech krocích výběru
- ◆ podrobné kontaktní formuláře (detailní interní informace o nenavrácených dotaznících, základní sociodemografické údaje o nerespondentech, typ sousedství a typ bytové jednotky, charakteristika okolí bydliště, záznamy o jednotlivých návštěvách, důvody neochoty odpovídat, popř. nemožnosti odpovědět atp.)
- ◆ sběr pomocných dat (externí informace z registrů a jiných administrativních záznamů a výzkumů)
- ◆ kontrola výrazných změn ve strategiích výzkumu v každé zemi a lišících se podmínkách, za kterých byla šetření realizována, a jejich dopady na aplikaci strategie zvyšování návratnosti

## 7 Analýza mezinárodních dat v komparativních výzkumech

Nenahraditelný teoretický přínos komparativních výzkumů je v jejich primární snaze nahradit proměnnou „země“ jinou vysvětlující proměnnou, která funguje na systémové úrovni. Tím, že se objasní například odlišné chování mužů a žen na jiném stupni, než na individuální úrovni odlišnosti pohlaví, je první přínos vysvětlující proměnné. Druhý tkví v tom, že pokud jsou vědci schopni vysvětlit odlišné chování mužů a žen nebo názory obyvatel na imigranty v zemi díky proměnné, která má systémový charakter, je tedy dílem sociálního prostředí města, země, národů (např. náboženský charakter země, specifické kulturní zvyky, zločinnost velkých měst, sociální kontrola malých vesnic), dokáží tím osvětlit chování různých národů zároveň. Díky společnému náboženskému charakteru zemí Norska, Dánska, Švédska, Velké Británie může být vysvětleno chování všech jejich obyvatel, všech Norů, Švédů, Dánů, Angličanů atp. Tím, že jsou komparativní výzkumy schopny nalézt tyto vysvětlující proměnné na systémové úrovni, mohou vysvětlovat chování různých národů termíny jedné a též teorie.

Analýzy komparativních výzkumů se tedy dostávají za hranice pouhého kategorizování různých zemí, jejich záměrem je nalézt vysvětlení pro seskupení podobných a odlišných zemí a zastoupit názvy zemí právě těmito vysvětlujícími proměnnými. Z tohoto důvodu je ale nezbytné, aby komparativní výzkumy do vytříbenějších statistických analýz poskytly nezávislou proměnnou země (společnosti) pro odhalování mezinárodních odlišností a mapování skrytých struktur v datech. Celý tento proces spěje k nahrazení unikátních jevů obecnými výroky a akumulaci velkého množství znalostí, ze kterých se systematicky vyvíjejí nové teorie. Každá vytvořená teorie se musí ale neustále ověřovat, testovat a modifikovat nebo zamítnout. Jako skvělý prostředek k celému tomuto procesu ověřování teorie se opět jeví promyšlené komparativní výzkumy, neboť poskytují širokým množstvím dat z kontinuálních šetření kvalitní empirické zázemí, které může sloužit k prověřování teorií v různých sociálních prostředích a tím zajistit jejich validitu. Tento cíl je ještě navíc umocněn možnostmi sekundárních analýz, které mohou s daty pracovat neomezeně, resp. nezávisle na prvotním záměru primárního výzkumu, a hledat nejrozumnější vysvětlující proměnné. Jako celek slouží pojednávané komparativní výzkumy k propojování teorie a empirie zároveň.

Než jsou komparativní výzkumy schopny se dostat k vytvořené teorii, musí projít celým empirickým procesem statistické analýzy dat. Základním problémem v analýzách komparativních výzkumů je velké množství dat. Na tento problém statistika reaguje aplikací různých multidimenzionálních technik, do kterých se standardně zařazuje faktorová analýza

(analýza hlavních komponent), diskriminační analýza, analýza latentních struktur, korespondenční analýza, shluková analýza, vícenásobná regresní analýza nebo multidimenzionální škálování. Jeden z tradičních způsobů redukce dimenzionality dat je užití pouze jedné multidimenzionální techniky, nejčastěji faktorové analýzy a metod, které výzkumníkům umožňují sledovat vztahy mezi proměnnými. Nicméně, pokud si badatelé takto přejí stanovit, které společnosti jsou si podobné či rozdílné skrze výsledky faktorové analýzy, ve studiích se opětovně objevují mnohé nejasnosti. Podle jednoho faktoru lze usoudit na podobnost zvolených zemí, podle jiného faktoru mohou být společnosti velice rozdílné. Další analytický problém vyrůstá ze závislosti výsledků na použitém druhu faktorové analýzy, například ve faktorové analýze může ortogonální varianta podat jiné výsledky než neortogonální technika v kontextu závislých (resp. nezávislých) faktorů [Maxwell, Pryor, Smith 2002: 22].

Jako mnohem lepší varianta zpracování velkého množství dat se vůči teoretickým cílům komparativních výzkumů jeví vhodná kombinace několika multidimenzionálních technik. Jako první statistický krok je ideální aplikace shlukové analýzy. Shluková analýza dokáže vytvářet kvalitní přehled v tak velkém množství informací. Disponuje schopností seskupit podobné jedince do jedné skupiny a oddělit je od skupin nestejných jedinců a tím skvěle slouží prvotním potřebám sekundárních analýz kvantitativních mezinárodních dat. Dokáže nastínit nové teoretické typologie a efektivně pracovat s nezávislou proměnnou země jako klasifikační proměnnou pro tvorbu shluků podobných jedinců. Identifikuje tak, zda má skutečně společnost jako kontrolní proměnná v analýze očekávatelný vliv na pozorované sociální jevy. Nejenže umožňuje shluková analýza vytvořit přehlednou strukturu velkého množství dat, které jsou pro komparativní výzkumy typické, dokáže navíc i zachytit první signifikantní proměnné, které by mohly mít na systémové úrovni žádoucí vliv a vysvětlující potenciál pro různé skupiny objektů (společností, zemí, národů, kultur, komunit, jedinců atp.). Pokud se odvozená typologie podobných a rozdílných objektů včetně nastínění prvních systémových proměnných zdá výzkumníkům rozumná, může shlukovací analýza poskytnout užitečný startovací bod pro další analýzy.

Následná aplikace ostatních multidimenzionálních technik (např. regresní analýzy) pak umožňuje v různých situacích prohloubení analýzy vztahů proměnných a skupin jedinců, nastíněných již shlukovou analýzou, a má schopnost nalézt proměnné, které na systémové úrovni fungují a ovlivňují individuální chování lidí, a vytvářet tak obecné výroky do vznikající teorie [příklad studie Wolfsona, Madjd-Sadjadi a Jamese 2004]. Tím pádem se



celý empirický rámec zkoumání uzavírá a komparativní výzkumy mají možnost naplnit své prvotní cíle, mít zásluhu na rozvíjejících se konzistentních a přesných sociálních teoriích, které budou empiricky testovatelné.

Neboť mezinárodní data je vždy nutno nějakým způsobem analyzovat, aby mohl komparativní výzkum dosáhnout výše zmíněných teoretických cílů, tato práce se bude nadále srovnávacím analýzám mezinárodních dat věnovat.<sup>14</sup> Není možné výkladově pokrýt použití všech multidimenzionálních technik v analýze mezinárodních dat v jejich návaznosti, obzvláště kvůli jejich konkrétnímu využití na specifická data, které je svázáno s jasně definovanými předpoklady každé jednotlivé metody a konkrétním výzkumným záměrem; jak již bylo výše řečeno, většina multidimenzionálních technik dokáže jít do větší hloubky analyzovaných vztahů. Tato práce se dále zaměří na první optimální krok ve statistické analýze velkého množství dat a to použití shlukovací techniky.

Shlukovací technika má jako jedna z nejpoužívanějších klasifikačních technik schopnost vytvořit první shrnutí nepřehledného množství různorodých dat. Dokáže se podívat na data a poradit výzkumníkovi, kterým směrem by bylo vhodné se v komparativních analýzách dále vydat a kterým ne. Navíc dokáže efektivně navrhnout novou teoretickou typologii, kterou v datech nalezne, a disponuje velkými klasifikačními možnostmi, což má neocenitelný význam pro jakoukoli vznikající teorii. Na druhou stranu se řadí spíše k explorativním metodám bez hlubšího kontextu, tudíž není ani zároveň vhodné její výstupy příliš interpretačně přeceňovat. Způsobu systematického využití shlukovacích technik na mezinárodní data se ještě stále odborná literatura příliš podrobně nevěnuje [Maxwell, Pryor, Smith 2002], jejich účelové a vhodné využití je však výzvou pro cíle komparativních výzkumů, z tohoto důvodu bude o shlukovacích metodách v této práci nadále pojednáváno.

V následujících kapitolách se tato práce nejprve pokusí o shrnutí teoretické podstaty shlukovací metody, což znamená vystihnout její původ v klasifikačních analýzách, historii jejího vývoje, teoretické vymezení její povahy a přínosů do analýzy mezinárodních dat a kategorizaci metod, které shlukovací technika standardně používá. Neboť prvotním záměrem nebylo detailně popsat všechny dostupné techniky shlukovací metody, v další části se tato práce řídí orientací sociálních věd a vyzdvihuje dvě nejdostupnější a nejpoužívanější shlukovací metody v analýzách mezinárodních sociálních dat, jmenovitě hierarchickou aglomerativní metodu a metodu K-průměrů. Navíc se k nim v koncepci práce připojuje i hlubší popis dvoukrokové shlukovací metody, dostupné ve statistickém programu SPSS,

---

<sup>14</sup> Některým teoretickým aspektům analýzy mezinárodních dat a požadavkům na jejich kvalitu jsme se již věnovali výše (viz. Kapitola 3, Kapitola 5).

kteřá reprezentuje pomyslný přechod mezi tradičními klastrovacími metodami a modifikovanými moderními přístupy a nabízí tak širokou škálu využití, a to obzvláště pro analýzu komparativních výzkumů, neboť je schopna se vyrovnat s velkým množstvím dat a proměnných. Všechny tři zmíněné shlukovací metody, o nichž bude dále pojednáváno, doplňují příklady jejich konkrétního využití na datech z mezinárodního výzkumu European Social Survey s názornými ukázkami výstupů dat.

V závěru práce jsou zmíněny některé varianty nejpokročilejších metod, ne tak snadno dostupných v uživatelských statistických programech. Tyto techniky jistě do budoucna slibují širší možnosti využití. Jejich algoritmy jsou již předem implementovány do technik tak, aby byly připraveny na velké množství proměnných i objektů mezinárodních výzkumů. Efektivně se vyrovnávají s chybami a omezeními tradičních metod a jsou přístupnější validizačním technikám výsledného řešení klastrování. Avšak jejich velkou nevýhodou je technický vzdor vůči běžným uživatelům. V podstatě ale vždy záleží na výzkumném zaměření samotného badatele, i méně pokročilé klasické techniky slibují kvalitní výsledky, jen je nutné sledovat jejich pravidla použití a nepřeceňovat výstupy explorativní shlukové metody.

## 8 Principy a povaha shlukové analýzy

### 8.1 Klasifikace, typologie a taxonomie

Vedle toho, že je *klasifikace* základní lidskou abstraktní činností, je také v nejširším smyslu slova pravděpodobně jednou z nejstarších vědeckých snah. Klasifikace hraje důležitou roli v rozvoji mnoha vědních disciplín a jejich odvětví. Není jen základem pro rozvoj jazyka a řeči, ale také pro matematiku, statistiku a obecně analýzu dat [Bailey 1994]. Jako názorný příklad poslouží biologická teorie evoluce, klasifikace chemických prvků či seskupování hvězd v astronomii. Klasifikace objektů (v tomto smyslu nejčastěji jedinců, resp. respondentů) do smysluplných podsouborů je ale také důležitou komponentou každého sociálně vědeckého výzkumu [Aldenderfer, Blashfield 1984; Landau, Everitt 2004]. Je základním procesem všech vědeckých praxí, neboť utváření klasifikačního systému znamená pokrok v rozvoji teorie uvnitř samotné vědy. Bez klasifikace nemůžeme dělat pokroky ani v přírodovědních disciplínách, ani v sociálněvědním výzkumu, i přesto ale není jádru tohoto procesu příliš dobře rozuměno. Klasifikace většinou vystupuje jako metodologický ekvivalent elektřiny, kterou dennodenně používáme, ale často ji považujeme za velice tajemnou a záhadnou [Bailey 1994].

Jako obecný termín značí klasifikace proces uspořádávání objektů do skupin nebo tříd na základě vlastní podobnosti. Statisticky řečeno, výzkumníci v klasifikaci požadují minimalizaci vnitroskupinové variance a maximalizaci meziskupinové variance. To znamená, že shromažďují skupiny entit tak, aby byly shluky těchto objektů navzájem co nejvíce odlišné a přitom byly jako samostatná jednotka vnitřně co nejvíce homogenní [Bailey 1994]. Takto definována může být klasifikace vnímána jako proces a zároveň jako výsledek.

Schopnost řadit podobné objekty do kategorií je možná na první pohled primitivní. Fakt, že klasifikace ale není lehký proces, dokazuje Bailey [1994] na příkladu představy 30 nožů, vidliček a lžic, které jsou dány k seřídění třem různým lidem. Jaké překvapení nás čeká, když se na konci setkáme se třemi různými výsledky, protože první člověk seřadil objekty podle velikosti, druhý podle materiálu, ze kterého jsou vyrobeny, a třetí podle druhu, kam patří, nože k nožům, vidličky k vidličkám a lžice ke lžicím. Následuje otázka, která klasifikace je lepší.

Bailey jasně ukazuje důležitost dimenze, podle které seskupujeme objekty, typů proměnných, na základě kterých klasifikujeme. Složitost seskupování na základě podobnosti objektů exponenciálně vzrůstá i s počtem objektů, které mají být klasifikovány, a počtem

dimenzí, na jejichž základě mají být dané objekty roztrženy. Základ úspěšné klasifikace tedy podle tohoto návodu leží ve schopnosti nalezení klíče nebo jinak řečeno v základním charakteru badatelova konceptu, podle kterého se dále klasifikuje. Bailey upozorňuje, že bohužel neexistuje přesný návod k nalezení tohoto klíče, ani upřesnění, zda je to otázka teoretické konstrukce, klasifikace či statistické analýzy.

Klasifikaci, která takto vyplývá ze sdružených objektů, lze rozdělit do dvou základních přístupů: typologie (teoretického přístupu) a taxonomie (empirického přístupu) [Bailey 1994: 6]. *Typologie* je jiný termín pro klasifikaci, ale odlišují ho dvě charakteristiky. Typologie je obecně multidimenzionální a teoretické povahy, funguje jako pojmenování teoretických typů. O typologii se často hovoří jako o kvalitativní klasifikaci, protože je většinou vytvářena bez potřeby kvantifikace či statistických analýz. Z tohoto úhlu pohledu je také více využívána v sociálních vědách. *Taxonomie* se od typologie odlišuje ve formě výsledků, neboť vyjadřuje klasifikaci empirických objektů. Její podstatou je nalezení reálných případů pro vytvoření teoretické typologie (opak teoretického přístupu), z tohoto důvodu se o ni hovoří jako o kvantitativní numerické taxonomii. Taxonomie je mnohdy hierarchická a vývojová a je častěji užívána v biologických vědách. Nejupřednostňovanější empirickou technikou v taxonomickém přístupu je shluková analýza.

Shluková analýza je obecné pojmenování rozsáhlého množství procedur, které jsou využívány k tomuto druhu empirické klasifikace, ale poněkud se liší od obecných klasifikačních analýz (např. diskriminační analýzy). V klasifikačních analýzách se pozorování seskupují do známého počtu predefinovaných skupin nebo populací, narozdíl od toho shluková analýza předem nezná ani počet skupin, ani skupiny samotné [Rencher 2002: 451]. Z hlediska statistiky, kde shluková analýza nachází své uplatnění, se shlukovací techniky řadí do multidimenzionálních technik (faktorová analýza, diskriminační analýza, korespondenční analýza, vícenásobná regresní analýza, multidimenzionální škálování atp.), které pracují již se složitějšími typy analýz a větším množstvím dat a proměnných. Obecně převažuje pro tento typ analýz termín *shlukovací techniky* (clustering techniques) nebo klastrovací techniky, některé vědní disciplíny používají pojmy seskupování souborů, numerická taxonomie, analýza shluků, explorační (průzkumová) analýza dat; někdy se dokonce klastrová analýza obecně označuje jako data mining - dolování informací z dat.<sup>15</sup>

---

<sup>15</sup> Data mining je označení analytických metod pro získávání skrytých a potencionálně užitečných informací z dat a vytváření statistických modelů, které se začaly masově rozvíjet v 90. letech 20. století; největší poptávka rostla v USA ze strany komerčních organizací, disponující velkými objemy dat, které vyžadovaly kvalitní materiály pro rozhodování o marketingových strategiích. [Šarmanová 2002]

Shlukové metody mají široké uplatnění v biologii, kognitivní a sociální psychologii, kriminologii, sociologii, antropologii, geologii, geografii, výzkumu trhu a ekonomii [Rencher 2002: 451]. Konkrétně lze pro ilustraci uvést studie popisu sociálních sítí a analýzy sítě přátelství v sociální psychologii, studie změn sociální struktury v sociologii, psychologické práce o klasifikaci mentálních onemocnění nebo kognitivních a intelektových schopností dětí, vytváření diagnostických manuálů a kategorizací nemocí v medicíně, vytváření typologií nástrojů a uměleckých předmětů v archeologii.

## 8.2 Historie klasifikačních klastrovacích technik

Po dlouhém období rozvoje teoretických klasifikačních procedur nastal příchod kvantitativních klasifikačních technik, jako je například zmíněná klastrová analýza, které částečně teoretický přístup nahradily; jejich primárním úkolem ale bylo klasifikační proces značně ulehčit a zdokonalit. Tento okamžik se udál ještě před postupným zaváděním počítačů do výzkumné oblasti. První klastrovací metody byly vymyšleny v období mezi 30. a 60. lety 20. století bez jakéhokoli přispění vyspělejší techniky. Svůj původ má tato kvantitativní metoda v psychologii, kde je spojována s pojmem modelová či systémová analýza (pattern analysis). Byla vytvořena psychology Zubinem (1938) a Tryonem (1939) a antropology Driverem a Kroeberem (1932). V té době popsal R. C. Tryon shlukovou analýzu následujícím způsobem: „*Shluková analýza je obecný logický postup formulovaný jako procedura, pomocí níž objektivně seskupujeme jedince do skupin na základě jejich podobností a rozdílností.*“ [podrobněji Bailey 1994: 34]

Až na konci 50. let McQuitty (1957) pokračoval v navrhování klastrovacího algoritmu, který se již bez počítačové techniky neobešel [podrobněji Bailey 1994: 11]. Následoval náhle prudký rozvoj tohoto statistického přístupu, jako první se rozrostly hierarchické aglomerativní metody, které primárně odhalily v procesu analýzy existenci jádra shluku, a rozvoj pokračoval zdokonalováním příslušných vzorců a formulí až ke konstrukci dalších pokročilejších shlukovacích metod [Bailey 1994: 11]. Hlavní stimul pro rozvoj klastrovacích metod byla v této době kniha *Principles of Numerical Taxonomy* publikovaná v roce 1963 biology Robertem Snokalem a Peterem Sneathem [podrobněji Aldenderfer, Blashfield 1984: 7].

Vývoj ale neustal, pozdější techniky jsou od brzkých 60. let až do současnosti většinou a exkluzivně počítačovými metodami. Všechny skupiny a podskupiny konkrétních přístupů se staly více komplexní a postoupily na vyšší úroveň. Divizivní hierarchické metody

se staly protiváhou vyvažující na vyšší úrovni hierarchické aglomerativní metody (viz. Kapitola 8.4). MacQueen publikoval v roce 1967 algoritmus K-průměrů (K-means) [podrobněji Bailey 1994: 60]. Navíc statistika začala čím dál více oceňovat kombinovatelnost a zároveň vyhraněnost jednotlivých přístupů, které tak mají správně zvoleným postupem rostoucí šanci zachytit provázanost jevů sociálního světa. V posledních letech 20. století došlo v souvislosti s některými novými objevy k renesanci zájmu o matematické modely neuronových sítí a jejich aplikaci a další moderní modifikace klasických shlukovacích metod [Řezanková, Húsek, Snášel 2009].

### **8.3 Představení shlukové analýzy**

Metody shlukové analýzy se vyvinuly z potřeby analyzovat informaci obsaženou v datech, která je generována množinou objektů, o jejíž struktuře není téměř nic známo. Strukturou se přitom myslí rozdělení objektů (případů, jedinců, v sociálních vědách respondentů) do určitého systému kategorií. Takovýmto roztríděním objektů do několika podsouborů následně vzniká klasifikace, jež vede k vytvoření systému tříd; přičemž algoritmy shlukové analýzy se mohou použít k měření jak starých klasických kategorizačních struktur, tak i k nalézání nových struktur. Tato její zmíněná vlastnost je obzvláště cenná v komparativních výzkumech, neboť svou klasifikační schopností dokáže vyvíjet nové teorie a zároveň ověřovat starší teorie, které musí být nějakým způsobem modifikovány.

Jestliže není známa kategorizace objektů, je cílem výzkumníka najít takovou kategorizační strukturu, jež je ve velké shodě s poměry v datech [Hendl 2004: 460]. Jinými slovy výzkumník hledá přirozené skupiny, které co nejpřesněji odrážejí informace zachycené v datech a dávají věcný smysl [Renchner 2002: 451]. Shlukovací metody jsou úspěšné především v situacích, kdy mají objekty (jedinci) tendenci seskupovat se právě do přirozených tříd, spíše než v případě náhodného rozmístění objektů v prostoru, do kterého by mohla být struktura dat vnucena.

Klastrová analýza je výborný popisný nástroj, který redukuje komplexitu velkého množství dat tím, že je klasifikuje do několika skupin. Tímto procesem umožňuje rozpoznat podobnosti objektů a seskupit podobné případy k sobě do jednotného shluku. Dokáže zachytit vztahy mezi zvolenými proměnnými a vytvořit tak komplexní pohled na sdružení objektů a zároveň i seskupení charakteristických proměnných, které by mohly mít systémový vliv na zkoumané sociální jevy a vysvětlit povahu seskupení daných objektů v komparativních

analýzách.

Její nevýhodou je fakt, že někdy nevyhovuje explanačním a predikčním výzkumným záměrům, které se v sociálních vědách vyskytují velice často. Občas také dochází k zvěčňování (reifikaci) teoretických jevů na empirické objekty, a přichází na řadu otázka, co je skutečně vykonstruované a co ne, zda není struktura do dat jen statisticky vnucena. Shluková analýza pak může začít postrádat svou teoretickou a konceptuální užitečnost. Nezřídka se shlukové analýze také vyčítá zachycení pouze statického charakteru dat, ale postrádající dynamické vlastnosti lze nahradit například analýzou kontinuálně sbíraných dat, což je v komparativních výzkumech velice běžný způsob sběru dat. Posledním problémem je skutečnost, že definice klastru je velice nejasná vzhledem ke konkrétnímu vybranému měření vzdálenosti objektů od klastru, které je shlukovací technikám vlastní. Tento nedostatek přesné definice klastru by se měl v některých případech stát překážkou výběru shlukovací metody do analýzy. Výzkumník musí totiž jasně rozumět jejímu použití a implicitním a explicitním požadavkům využívaných měření vzdáleností objektů [Gower 1967].

Aldenderfer a Blashfield [1984: 9] vymezují nejzákladnější teoretické cíle shlukové analýzy: rozvíjet typologii nebo klasifikaci užitečných sociálních jevů, konstruovat hypotézy na základě explorační dat a testovat tyto hypotézy. Vedle toho lze pro shlukovou analýzu stanovit i konkrétnější empirické druhy úloh, které by měla naplňovat. Jejím cílem by mělo být nalezení předem definovaného množství shluků, nalezení množství shluků, jejichž počet není předem specifikován, a nakonec vytvoření hierarchického stromu (tzn. vytvořit posloupnost spojování objektů k sobě) [Hendl 2004: 461]. Pokud shluky existují, úkolem shlukové analýzy je skupiny charakterizovat a interpretovat, a navíc pokud jsou do analýzy zahrnuty další objekty, shluková analýza by měla být schopna je zařadit do již definovaných klastrů [Šarmanová 2002].

Shluková analýza netvoří ucelenou teorii, ale je to řada metod založených na různých principech. Například kromě objektů se mohou podrobit shlukové analýze i vlastní proměnné (tzv. R-technika shlukování). Shluková analýza se provádí zpravidla na množině objektů, kde je každý objekt popsán řadou znaků (proměnných), takový postup se označuje jako Q-technika shlukování [Hendl 2004: 424]. Na druhé straně R-technika shlukování vychází z analýzy množiny proměnných, charakterizovaných prostřednictvím objektů. Podobný duální přístup se uplatňuje i v dalších multidimenzionálních metodách, například ve faktorové analýze. Shlukování proměnných klastrovacími metodami má společné znaky s klasickou faktorovou analýzou, obě metody se snaží prozkoumat vztahovou strukturu mezi proměnnými

a redukovat množství proměnných charakterizující objekty, ačkoli využívají úplně rozdílné techniky [Hendl 2004: 461; DiStefano, Kamphaus 2006: 780].<sup>16</sup>

Shlukovací techniky je dále možné vymezit podle různých kritérií, např. zda je na začátku určen počet shluků nebo se má v průběhu řešení nalézt jejich optimální počet, a nebo podle výsledné struktury skupin objektů (počet objektů ve shluku, tvar klastrů, počet prezentovaných klastrů, stupeň jejich překrytí, pokrytí dat). V procesu shlukovací analýzy se používají různé míry vzdáleností (metriky) objektů, resp. míry podobností a nepodobností, také různé algoritmy (postupy) shlukovacích metod při zpracování. Výsledky klastrové analýzy navíc ovlivňuje různorodost řešených problémů, požadovaný typ výsledků, velké množství dat, závislostní struktura zahrnutých proměnných atp.

Pro řešení základních otázek klastrové analýzy neexistují žádná systematická pravidla. Z důvodů potenciálních různých výsledků klastrovacích metod je důležité dbát na základní pravidla jejich použití a uvědomovat si, že jemně odlišné výsledky analýz (použitých na totožných datech) jsou většinou způsobené vlastními charakteristikami a postupy technik samotných, a velmi odlišné výsledky analýz mohou být způsobené vnitřní strukturou odchýlených dat. Klastrová analýza navíc přirozeně vkládá do vzorku hierarchickou strukturu, a i když je to legitimní požadavek pro typologické cíle, může být poněkud nevhodná v jiných situacích, ve kterých by měla být zastoupena jinými multidimenzionálními statistickými technikami [Gower 1967].

Je dobré upozornit na to, že výzkumník musí mít stále na paměti, že klastrovací metody jsou spíše heuristické, nezaručují nejlepší řešení a jsou založené na zkušenostech a praxi výzkumníků, na rozdíl od faktorové analýzy nebo regresní analýzy, jejichž principy spočívají v rozsáhlém a odborném statistickém základu [Aldenderfer, Blashfield 1984: 14]. Ačkoli i klastrovací techniky mají k dispozici signifikantní matematické nástroje, je nutné si uvědomit, že podstata těchto metod je velice jednoduchá a prostá. Pokud si toto výzkumník připustí, vyhne se chybám v interpretaci klastrového řešení. Klastrovací metody nejsou často nic víc než přijatelné (ale také věrohodné) pravidlo pro shlukování objektů. Z tohoto důvodu může být zajímavější a podnětnější soustředit se na vnitřní kohezivitu shluku než na jejich vnější vzájemné odlišnosti [Everitt 1979: 177]. Důležitější je to, co shluk tvoří, které proměnné ho dělají jedinečným.

I přes popularitu shlukovací analýzy a rychlost jejího rozvoje není této technice stále příliš dobře rozuměno ve srovnání s ostatními multidimenzionálními technikami jako je

---

<sup>16</sup> O využití faktorové analýzy jako Q techniky více Everitt [1979: 170] a Aldenderfer, Blashfield [1984: 49].



faktorová analýza, diskriminační analýza, analýza latentních struktur nebo multidimenzionální škálování. Sociálněvědní literatura věnující se klastrování se často potýká se zarážející a protichůdnou terminologií, metodami a preferovanými přístupy shlukových metod. Tato rozmanitost podání vytváří složitou nekonzistentní zmrš, do které je někdy velice náročné proniknout. Adekvátně pojednávat o každé klastrovací metodě individuálně je tedy téměř nemožné.

Klasické shlukovací metody jsou přístupné v softwarových nástrojích pro analýzu vícerozměrných dat. Mezi relativně dostupné uživatelské statistické programy, nejzmiňovanější v odborné literatuře, patří Statistica, S-Plus, SPSS, SYSTAT, SAS, volně dostupný program R a klastrovací program LOICZView či ClustanGraphics (publikovaný softwarový balíček Clustan, Ltd. SPSS).

#### **8.4 Kategorizace klasických metod shlukové analýzy**

Shluková analýza je společný název pro celou řadu exploračních statistických metod, jejichž cílem je využití informací z analýzy vícerozměrných dat k rozřídění množiny heterogenních objektů do několika relativně homogenních podsouborů, označených jako shluky (clustery). Objekty (pozorování, jedinci, respondenti atp.) uvnitř shluků si mají být co nejvíce podobné a objekty patřící do různých shluků co nejvíce rozdílné [Everitt 1979: 169; Rencher 2002: 451]. Pro objekty platí základní pravidla kompletnosti a vzájemné exkluzivity: každý objekt musí nalézt svou příslušnou skupinu a žádný objekt nesmí být členem dvou skupin zároveň, musí patřit výlučně pouze do jedné [Bailey 1994: 3]. Druhým úkolem klastrové analýzy je vytvořit hierarchii těchto shluků objektů.<sup>17</sup>

Základní přístup shlukové analýzy je takový, že každý objekt je jednoznačně zařazen do jednoho shluku. Toto shlukování se nazývá pevné a disjunktní (hard clustering). Někdy může být objekt zařazen do dvou shluků a výsledkem jsou tedy překrývající se shluky (overlapping clustering, soft clustering, někdy clumping). Obecně se klasické shlukové metody vyhýbají vzniku překrývajících se shluků a pro jejich identifikaci byly navrženy speciální metody překrývání klastrů. Pokud nejsou pro objekt k dispozici konkrétní hodnoty jejich znaků (proměnných), podle kterých jsou zařazovány do shluků, jde namísto pevných

---

<sup>17</sup> Jak hledaný shluk vypadá? Má svou hustotu jako nahuštěný roj objektů v prostoru, elipsoidní či hypersferický (kulovitý) tvar; rozptýl objektů v prostoru, které k sobě přilehají blízko středu a uvolňují se směrem ven; dimenzionalitu v podobě pole okolo shluku a izolaci jako míru s jakou se klastry překrývají [Aldenderfer, Blashfield 1984: 34]. Shluky jsou stabilní, mají tendenci reprezentovat přirozené skupiny v datech a nezmizí, pokud jsou objekty reorganizovány nebo několik objektů odstraněno.

dat o tzv. *fuzzy data*. Až teprve kombinace fuzzy dat a překrývajících shluků vytváří prostředí pro moderní fuzzy shlukovou analýzu, kdy je ke každému objektu nejprve přiřazena míra příslušnosti ke každému ze shluků, na základě níž je teprve provedeno přiřazení objektu do klastru [Řezanková, Húsek, Snášel 2009: 17].

Danou množinu objektů lze často roztrždit mnoha rozdílnými způsoby (metodami), přičemž každá získaná struktura má svůj význam. Každá klasifikace může odrážet jiné aspekty dat [Hendl 2004: 464]. Z hlediska analýzy dat je vstupem pro shlukování vždy datová matice a výstupem je identifikace shluků, které mohou být různých typů. V procesu klastrování pak hrají roli kritéria shlukování, stanovení počtu shluků a interpretace výsledků [Řezanková, Húsek, Snášel 2009: 24]. Podle všech těchto aspektů lze vytušit značně nejednotnou klasifikaci shlukovacích metod. Navíc byla v poslední době navržena řada nových algoritmů, některé modifikují tradiční metody shlukové analýzy, jiné se vydávají novými směry. Následující odstavce několik klasifikačních pokusů shrnují.

V terminologii klasifikačních metod se primárně rozlišuje:

- ◆ *učení s učitelem* (supervised learning)
- ◆ *učení bez učitele* (unsupervised learning, machine learning)

V metodě učení s učitelem obsahuje vstupní datový soubor řadu informací o příslušnosti objektů do známých skupin. Jejím cílem je vytvořit model, na základě něhož by mohly být objekty bez známé příslušnosti zařazovány do daných skupin [Řezanková, Húsek, Snášel 2009: 23]. Nejčastějším příkladem tohoto přístupu je klasifikování nových dat na základě již existující typologie (např. diskriminační analýza).

V přístupu učení bez učitele není předem známa příslušnost žádného z objektů a obvykle není znám ani počet skupin (musí se vybrat proměnné, jejich relativní váhy a finální počet klastrů). Jeho cílem je klasifikovat všechny objekty zahrnuté do analýzy, tento postup je často označován jako shlukování (resp. analýza vyvíjející novou typologii). Oba přístupy lze použít ke klasifikaci objektů, i když se někdy termín klasifikace používá v užším slova smyslu pro metody vycházející z principu učení s učitelem [Řezanková, Húsek, Snášel 2009: 23].

Podle intenzity zásahu výzkumníka do analýzy se rozlišuje *subjektivní* a *objektivní* klastrování. V případě subjektivního shlukování výzkumník provádí analýzu na konkrétním vybraném souboru proměnných a sám si rozhoduje o úrovni vstupu objektu do klastru. Na druhé straně objektivní klastrování zjišťuje výpočty nejvýhodnější kombinaci proměnných, které nejlépe zařazují objekty do shluků, a jediné kritérium vstupu objektu do shluku je nejmenší vzdálenost tohoto objektu od příslušného shluku (resp. princip největší

podobnosti) v každém kroku analýzy [Bailey 1994: 55].

Další klasifikace vychází z matice vzdáleností, jejímiž prvky jsou hodnoty vypočtené pro všechny dvojice objektů na základě míry nepodobnosti, která vychází z datové matice. Pro vyjádření vztahu dvou objektů se ve shlukovací metodě klasicky používají především míry vzdáleností, které říkají nakolik jsou si dané objekty v prostoru vzdálené, tedy nepodobné. Reprezentují tímto způsobem podobnost objektů, neboť čím blíže si objekty v prostoru jsou, tím podobnější se stávají.<sup>18</sup> Podle matice vzdáleností rozlišuje Řezanková, Húsek a Snášel [2009: 29]:

1. *metody založené na vzdálenostech* (distance-space methods), které v podstatě reprezentují hierarchické metody shlukování
2. *metody vektorového prostoru* (vector-space methods), které se naopak přibližují nehierarchickým metodám a dále se dělí na:
  - ◆ *inkrementální* – postupně procházejí jednotlivé objekty a přímo je zařazují do shluků, klasifikují převážně velké datové soubory a závisí na pořadí objektů
  - ◆ *optimalizační* - vycházejí z počátečních shluků, někdy náhodných, a iterativně zlepšují výsledky zařazování objektů do klastrů

Jako tradiční klasifikace shlukových metod vystupuje rozdělení *hierarchických a nehierarchických technik* podle výsledku shlukování. Ani tato klasifikace není jednoznačná, neboť mnohé odborné publikace pojmenovávají nehierarchické metody jinak, např. je označují jako disjunktí techniky nebo nejčastěji metody rozkladu (iterative partitioning methods, flat methods).

1. *Metody rozkladu* jsou založené na zařazování objektů do předem stanoveného počtu shluků, při tomto typu shlukování je tedy vytvářen konkrétní počet shluků. Přiřazení ke shlukům je buď jednoznačné, příkladem je známá metoda K-průměrů (K-means), nebo se počítá míra příslušnosti jednotlivých objektů ke shlukům, která je zajištělná pomocí fuzzy shlukové analýzy [Řezanková, Húsek a Snášel 2009: 84]. Metody rozkladu se dále dělí na:
  - ◆ *optimalizační*
  - ◆ *analýzu modů*
2. *Hierarchické metody* shlukování fungují jako postupné rozkládání, které na jedné straně začíná úplným rozkladem, v němž každý objekt tvoří jednoprvkový shluk, a

---

<sup>18</sup> Zmíněný prostor vzniká na základě kombinace dimenzí, které vytváří vybrané klasifikační proměnné [Bailey 1994: 9]. Pokud jsou ke klasifikaci vybrány 4 proměnné, vytvoří tak čtyř-dimenzionální prostor.

postupně se tyto shluky spojují po dvojicích až do vytvoření jednoho společného velkého shluku. Na druhé straně začínají s jedním shlukem obsahujícím všechny objekty a postupně se rozkládajícím až do nejmenších jednotlivých objektů jako jednoprvkových shluků. Podle směru postupu rozkladu při hierarchickém shlukování se dále metody hierarchického shlukování dělí na aglomerativní a divizní. Hierarchická metoda nikdy nevytváří konečný počet shluků, toto konečné rozhodnutí přísluší výzkumníkovi.

- ◆ *aglomerativní přístup* (Forward clustering, „bottom up“, **AGNES** **AG**lomerative **NE**Sting) – aglomerativní přístup značí směr hierarchického shlukování, kdy je každý objekt nejprve samostatným shlukem a postupně se na základě analýzy podobnosti po dvojicích tyto shluky spojují od nejvíce podobných shluků k nejméně podobným až do výsledného jednoho shluku. Na základě postupných kroků shlukování lze tuto metodu nazvat sekvenční. Zvláště v sociálních vědách se používají aglomerativní přístupy mnohem častěji [Rencher 2002: 455].
- ◆ *divizní přístup* (Backward clustering, „top down“, **DIANA** **DI**visive **AN**alysis) – divizní přístup hierarchického shlukování operuje s předpokladem, že na začátku všechny objekty tvoří jeden shluk a ten je opakovaně rozdělován na základě analýzy nepodobnosti (nejdříve na dva samostatné soubory a pak dále) až do stavu, v němž je každý objekt samostatným shlukem (resp. dokud není splněno kritérium, které ukončí analýzu, například předem definovaný počet kroků). Tato metoda má spíše simultánní charakter, vše se děje v jednom kroku. Nalezení optimálního rozkladu množiny je velice náročné a proveditelné jen pro malý počet objektů, divizivní přístup se tak používá obzvláště při preferování vzniku velkých klastrů [Rencher 2002: 479].<sup>19</sup>

---

<sup>19</sup> Divizní metoda se ještě dále může dělit podle počtu klasifikačních proměnných na monotetickou a polytetickou variantu [více Aldenderfer, Blashfield 1984: 50].

## 9 Vlastní využití klastrovacích technik na datech European Social Survey

V předešlých kapitolách jsme se pokusili nastínit na konkrétních příkladech užitečnost a důležitost využití shlukovacích metod v rámci sekundárních analýz komparativních výzkumů. Jak již bylo zmíněno, prvotním cílem této práce není detailně popsat využití všech technik shlukovacích metod. Nadále se budeme soustředit na dvě nejdostupnější a nejpoužívanější shlukovací metody v komparativních sociálních výzkumech, na hierarchickou aglomerativní metodu a metodu K-průměrů, a pokusíme se na jejich využití v mezinárodních výzkumech podívat nepatrně kritickým pohledem, neboť jejich výběr pro analýzu komparativních dat není vždy vhodný, a vzájemně porovnáme jejich výstupy.

Navíc připojíme popis analýzy pokročilejší dvoukrokové shlukovací metody, dobře dostupné ve statistickém programu SPSS, jako příklad modernější již modifikované hierarchické metody, která svým přístupem zacházení s daty připomíná statisticky vylepšené moderní shlukovací techniky, které se efektivně vyrovnávají s velkým množstvím dat a různorodými proměnnými. Již tato verze (dvoukrokové) shlukové analýzy se znatelně lépe výpočetně vypořádává s velkým množstvím dat a proměnných než například klasická hierarchická metoda. Z tohoto důvodu je také zařazena do koncepce této práce. Je dostupnější v uživatelském statistickém programu SPSS, lépe než zmíněné verze moderních vícerozměrných shlukovacích technik, a navíc svými pokročilejšími možnostmi výstupů demonstuje, jaké zlepšení lze v analýze mezinárodních dat od moderních metod očekávat.

Všechny tři vybrané shlukové metody, jež jsou dále detailněji rozpracované, jsou průběžně doplňovány příklady jejich konkrétního využití na datech z mezinárodního výzkumu European Social Survey s názornými ukázkami výstupů dat. Pro ilustraci výstupů jednotlivých shlukovacích metod autorka do konceptu práce zvolila data z první vlny European Social Survey, realizované mezi lety 2002 – 2003 [Jowell, Central Co-ordinating Team 2002/2003]. První vlny výzkumu se zúčastnilo následujících 22 zemí:

Austria	Hungary	Portugal
Belgium	Ireland	Slovenia
Czech Republic	Israel	Spain
Denmark	Italy	Sweden
Finland	Luxembourg	Switzerland
France	Netherlands	United Kingdom
Germany	Norway	
Greece	Poland	

Výzkum proběhl v České republice v listopadu a prosinci v roce 2002 pod záštitou Sociologického ústavu Akademie věd ČR. Terénní sběr dat měla na starosti agentura STEM s.r.o. (viz. Příloha 1). Obecně byl výzkum proveden podle standardních měřítek European Social Survey náhodným výběrem na základních principech pokrytí celé populace, vysoké návratnosti a žádné náhrady nekontaktovaných respondentů [ESS 2003]. V České Republice byl v první vlně výzkumu uskutečněn dvoukrokový stratifikovaný výběr, na jehož podkladě bylo vybráno nejprve 100 velkých měst pokrývajících 52 % populace následně rozdělených do menších tříd (pak byly vybrány domácnosti na základě SIPO a v nich jednotlivci na základě Kishovy mřížky), zbytek tvořily menší lokality rozkládané do dalších skupin pokrývajících 48 % obyvatel (podle Censu 2001) [ESS 2002b].

Minimální počet respondentů byl kvůli různým designovým váhám pro všechny země stanoven na hodnotě 2000 respondentů a 1000 respondentů v menších zemích pod 2 miliony obyvatel. V České republice byl tento minimální počet stanoven na hodnotě 2100 respondentů a při designové váze ( $n_{\text{eff}}$ ) 1,63 bylo ve výsledku náhodně vybráno 3300 domácností k provedení dotazníkového šetření. Bohužel, v prvním kole ESS byla v ČR míra návratnosti dotazníků poměrně nízká (43,3 %), z tohoto důvodu bylo ve skutečnosti provedeno jen 1360 rozhovorů. Je vhodné také podotknout, že v prvním kole ESS došlo k úmyslnému nadhodnocení minoritních skupin (východní Německo v Německu, arabská populace v Izraeli atp.) kvůli očekávané nižší návratnosti a horší kontaktovatelnosti členů různých subskupin. Datový soubor má v kompletní podobě 561 proměnných a 42 359 respondentů.

Strukturovaný dotazník distribuovaný v České Republice formou PAPI byl klasicky tvořen hlavním modulem a dvěma rotačními moduly. Při výběru klasifikačních proměnných do shlukové analýzy se autorka převážně soustředí na rotační modul „Imigrace a postoje“ (sekce D v dotazníku), poskytující širokou škálu spojitých i kategoriálních proměnných pro potřeby analýzy. Modul vytváří položky, které mají za úkol porozumět způsobu, jakým je v oblasti imigrantství formováno veřejné mínění a jak je spojeno s postoji k minoritám a různým subkulturám ve společnosti. Rotační modul má tři dimenze: tzv. škálu imigrace, která má za úkol odhadnout rozsah současné a minulé migrace, a jak ji lze srovnat s migrací v jiných evropských zemích; dále profily migrantů, kteří žádají o azyl nebo jsou prozatím uprchlíky; a nakonec míru diskriminace, které jsou imigranti a jiné subskupiny vystaveny. V neposlední řadě je cílem rotačního modulu analyzovat veřejné vnímání dopadu imigrace na ekonomiku státu a tržní prostředí, národní kulturu, chudobu, kriminalitu a státní výdaje.

Konkrétní klasifikační proměnné i výběr respondentů se volil pro daný typ shlukové

analýzy podle specifických požadavků konkrétní shlukovací techniky, které jsou u každé metody uvedeny, proto se v popisu každé techniky vyskytují grafické a numerické výstupy pro jiné proměnné. Výběr respondentů byl vždy totožný, do analýzy dat jsme zařadili všechny *občany* klasifikovaných zemí a vyřadili všechny respondenty, kteří nejsou občany daných zemí; validní (platný) počet respondentů, se kterým shlukovací analýzy dále pracují, je ale u každé metody odlišný. I přesto je zde patrná snaha dodržet komplexní charakter výstupů v rámci jedné metody, což je pro pochopení analytického procesu různých technik nejdůležitější.

Zvolené příklady nemají plnit funkci sekundární analýzy s jejími interpretačními požadavky. Každý daný výstup disponuje interpretačním potenciálem vzhledem k výběru proměnných z konkrétní oblasti imigrantství výzkumu ESS (2002), které společně s množstvím sebraných mezinárodních dat poskytují kvalitní podklad pro jakýkoli úzce zaměřený výzkum. Všechny uvedené výstupy by jistě v jiném druhu práce tuto funkci dokázaly naplnit, ale v této konkrétní studii slouží jako názorná pomůcka při ilustrování analytických schopností klastrovacích metod v případě rozsáhlých dat mezinárodního výzkumu. Dokládají obecná fakta obsažená v textu, slouží jako příklady konkrétních provedení shlukovacích metod a pomáhají lépe pochopit relativně technický popis metod tím, že podávají převážně grafické výsledky analýz sociálních jevů.

Obzvláště v tomto kontextu nelze od uvedených výstupů dat očekávat potvrzení teoretických cílů komparativních výzkumů a to nalezení vysvětlujících proměnných, které na vyšší systémové úrovni dokáží objasnit vznik jednotlivých shluků podobných zemí, které jsou z výstupů patrné. To by bylo úkolem právě takových interpretačních záměrů, které tato práce nemá. Naopak dokládané grafy jsou ukázkou velkého exploračního potenciálu shlukovacích metod a ilustrací prvotního přehledu, který shlukovací techniky ve velkém množství mezinárodních dat dokáží poskytnout (popř. prvním zachycením zajímavých souvislostí v datech), který si ale také v závislosti na konkrétních výzkumných cílech žádá další hlubší prozkoumání jiným druhem multidimenzionální analýzy.

### *1. Hierarchická aglomerativní metoda*

Z celého souboru jsme vyseparovali pouze ty respondenty z první vlny výzkumu ESS (2002), kteří *jsou občany* klasifikovaných zemí. Výsledný vzorek respondentů nabyl velikosti 40 646 respondentů, z nichž bylo do hierarchické shlukovací metody zařazeno 38 814 validních případů z 22 zemí. Kvůli požadavkům hierarchické aglomerativní metody na malou velikost vzorku (klasicky  $< 250$ ) jsme použítá data navíc do analýzy agregovali podle

proměnné země (country), tudíž se dále v analýze klastují jednotlivé země a ne jedinci. Hierarchická metoda tak zpracovává pouze 22 objektů (podle 22 zemí).<sup>20</sup>

Ke klastrové analýze byly vybrány dvě ordinální proměnné, co nejméně zkorelované - *důvěra v policii* a *ochota přijmout imigranta jiné rasy do rodiny* - mapované 10-ti bodovou škálou (vyšší body na škále znamenají u důvěry v policii větší důvěru, u ochoty přijmout přistěhovalce jiné rasy do rodiny značí vyšší skóre velkou neochotu). Kvůli značnému rozsahu škály se tyto ordinální proměnné dále při výpočtech považovaly za spojité. Znění otázek v češtině je následující:

„Řekněte mi prosím s použitím stupnice 0 – 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas (Policie). 0 znamená, že instituci vůbec nedůvěřujete, 10 znamená, že ji naprosto důvěřujete.“ (*důvěra v policii*)

„A nyní se zamyslete nad lidmi z jiných zemí, kteří přesídlili do České republiky a kteří jsou jiné rasy nebo etnické skupiny než většina domácího obyvatelstva. Do jaké míry by Vám vadilo, kdyby někdo takový uzavřel sňatek s Vaším blízkým příbuzným?“ (*ochota přijmout přistěhovalce jiné rasy do rodiny*)

Standardně se používá u všech aglomerativních metod vzdálenost Squared Euclidean distance z důvodu srovnatelnosti jednotlivých technik, které ukazují i přes rozdílné procesy shlukování ve výsledku podobná řešení, což ukazuje na neodchýlená kvalitní analyzovaná data. Základní vzdálenosti objektů (Euklidovská, Čebyševova, Manhattanská a Minkowského metrika) se ve výstupech neudávají, neboť produkují zcela totožné výsledky včetně celého hierarchického postupu shlukování a díky svým nedostatkům nejsou tak dobře srovnatelné ve všech použitých aglomerativních metodách.

## 2. Metoda K-průměrů

Do analýzy jsme opět zvolili *všechny občany* klasifikovaných zemí (40 646, jako validních se však v průběhu klastrování zařadilo jen 22 000 respondentů z 22 zemí). Tato metoda již umožňuje větší velikost vzorku (klasicky  $> 200$ ), ale většinou je vhodné nepřesáhnout počet 1000. Z tohoto důvodu jsou i pro tuto metodu použita data opět agregovaná podle proměnné země (country) a ve shlukování prostřednictvím metody K-průměrů figuruje jen 22 objektů.

Podle požadavků metody K-průměrů na nezávislost proměnných je vhodné do analýzy

---

<sup>20</sup> Agregací se dále myslí skutečnost, že jedinci jsou v analýze utříděni do zemí, ke kterým náleží. Do analýzy pak jako samostatné objekty vstupují jen tyto země, které jsou shlukovacími metodami dále klastrovány. Země vstupují do shlukovací analýzy navíc jen v podobě průměrných hodnot vybraných proměnných, které jsou vytvořeny pro každou zemi zvlášť.



opět zvolit co nejméně zkorelované proměnné, vybrali jsme dvě proměnné: *důvěra v policii* a *ochota přijmout imigranta jiné rasy do rodiny* (proměnné jsou totožné jako v hierarchické metodě). Obě proměnné jsou opět ordinální s 10-ti bodovou škálou, tudíž se považují v analýze za spojité (vyšší body na škále opět znamenají u důvěry v policii větší důvěru, u ochoty přijmout přistěhovalce jiné rasy značí vyšší skóre velkou neochotu). Do analýzy jsme zvolili v počátku 4 klastry (podle výsledků předešlých hierarchických metod).

### 3. Dvoustupňová shluková analýza

Do dvoustupňové shlukové analýzy jsme opět vybrali *všechny občany* klasifikovaných zemí. Z počátečního počtu 40 646 respondentů jich zbylo 33 509 validních zařazených do analýzy z 22 zemí. Při použití této metody se tentokrát postupovalo jiným způsobem a data se do analýzy neagregovala podle proměnné země, dvoustupňová shluková analýza již výpočetně zvládá tak velké množství jedinců v analýze. Z tohoto důvodu jsou do finálních klastrů zařazeni všichni jedinci ve výzkumu bez ohledu na jejich příslušnost k dané zemi. Tento krok se v analýze podnikl primárně proto, aby se mohl zdokumentovat vliv proměnné země separovaně od ostatních klasifikačních proměnných, namísto vytvoření klastrů zemí. Jako míra vzdálenosti shluků a objektů se používala Věrohodnost (Log-likelihood).

Tato metoda také umožňuje do analýzy zařadit velké množství proměnných, proto jsme se rozhodli pro následující kategoriální proměnné: *země, pohlaví, přátelé z jiné země* (Máte nějaké přátele, kteří se do České republiky přistěhovali z jiné země?), *pracovní kolegové z jiné země* (Máte v práci nějaké kolegy, kteří se do České republiky přistěhovali z jiné země?), *povolit vstup do země lidem stejné rasy a jiné rasy* (Do jaké míry by podle Vás měla Česká republika dovolit, aby sem přicházeli a žili zde lidé stejné rasy nebo stejného etnika/jiné rasy nebo etnické skupiny jako je většina českých lidí?), *lidem z bohatších zemí a chudších zemí* (Do jaké míry by podle Vás měla Česká republika dovolit, aby sem přicházeli a žili zde lidé z bohatších evropských zemí/z chudších evropských zemí?).

Jako spojité proměnné byly do analýzy zvoleny: *ochota přijmout do rodiny přistěhovalce jiné rasy* (A nyní se zamyslete nad lidmi z jiných zemí, kteří přesídlili do České republiky a kteří jsou jiné rasy nebo etnické skupiny než většina domácího obyvatelstva. Do jaké míry by Vám vadilo, kdyby někdo takový uzavřel sňatek s Vaším blízkým příbuzným?), *názory na ovlivnění ekonomiky země, zločinnosti v zemi a obohacení kultury imigrantstvem* (Řekl(a) byste, že je pro českou ekonomiku obecně špatné nebo dobré, že sem přicházejí žít lidé z jiných zemí? Řekla(a) byste, že lidé z jiných zemí, kteří sem přicházejí žít, většinou narušují nebo obohacují kulturu České republiky? Způsobují lidé z jiných zemí, kteří

sem přicházejí žít, v České republice zhoršování nebo zlepšování problémů spojených se zločinností?). U spojitě proměnné ochoty přijmout do rodiny přistěhovalce jiné rasy mají vyšší body na 10-ti bodové škále záporný význam, značí tedy neochotu; naproti tomu na 10-ti bodové škále hodnocení vlivu imigrantství na ekonomiku, kulturní rozvoj a zločinnost v zemi znamenají vyšší body kladné hodnocení tohoto vlivu (tzn. zlepšuje ekonomiku, obohacuje kulturu, pomáhá s problémy se zločinností).

Všechny uvedené metody a jejich výstupy byly zpracovány statistickým programem SPSS.

## 9.1 Hierarchická aglomerativní metoda

Jak již bylo zmíněno, princip aglomerativní shlukovací metody tkví v tom, že na začátku je každý objekt jednorvkovým shlukem. Postupně se po dvojicích tyto shluky na základě analýzy podobnosti spojují až k vytvoření jednoho konečného shluku. Hierarchický klastrovací algoritmus je v tomto významu nejen sekvenčním krokovým procesem, ale také ireverzibilním [Rencher 2002: 455]. To znamená, že žádné dva objekty spojené dohromady v jednom klastru od sebe nemohou být později v analýze odděleny, objekty nemohou být přemístěny a tím pádem ani případné chyby nemohou být později v analýze napraveny. Sneath a Sokal (1973) je označují akronymem **SAHN** – Sequential (sekvenční), Agglomerative (aglomerativní), Hierarchical (hierarchické), Nonoverlapping (nepřekrývající se) [podrobněji Bailey 1994: 48]. Hierarchické aglomerativní metody jsou nejpoužívanější shlukovací technikou v sociálních vědách [Bailey 1994: 48] a zatím také nejčastěji užívaným typem shlukovacích metod v komparativních výzkumech.

Cílem hierarchického klastrování není na konci analýzy dat volit optimální počet klastrů, více se soustředí na vytvoření kompletní stromové struktury; především tak vizuálně demonstruje celý proces postupného shlukování objektů [Everitt 1979: 172]. Hierarchické klastrování je vhodné pro menší vzorky (typicky < 250 objektů). Pokud je počet objektů příliš velký, algoritmus se dostává k výslednému řešení velice pomalu (tzn. doslova zahltní počítač, popřípadě se k němu nedostane vůbec). Z tohoto důvodu není hierarchická metoda vhodná pro analýzu velkého množství mezinárodních dat komparativních výzkumů, ale je ideální pro vytvoření představy možnosti shlukování na vybraném menším vzorku jedinců.

Její největší přínosy se v této oblasti jeví při použití hierarchické aglomerativní metody na shlukování proměnných, tedy ve formě R techniky. Zde má široké možnosti

uplatnění obzvláště při testování ekvivalence mezinárodních komparativních výzkumů. Jako konkrétní příklad může sloužit studie Billieta a Welkenhuysen-Gybelse [2002]. V rámci ověřování ekvivalence položek použitých v komparativním výzkumu ESS použili Billiet a Welkenhuysen-Gybels právě hierarchickou klastrovací analýzu. Pro názornou ukázkou si vybrali šest položek z oblasti imigrantství a zkoumali, zda položky sledují stejný teoretický jev ve všech účastnících se zemích první vlny sběru dat ESS (2002), a pokud ne, zda lze alespoň předpokládat společnou dílčí strukturu. K určení těchto skupin identických položek používali autoři hierarchickou Wardovu metodu, aby mohli zredukovat počáteční velkou skupinu proměnných (58 proměnných v sekci D). Jako jiný příklad může sloužit uplatňování shlukové hierarchické analýzy v replikačních studiích, které mají za úkol opětovně měřit validitu již navržených škál měření sociálních či psychologických jevů [více studie Worrella et al. 2006].

Celý proces aglomerativního shlukování v sobě zahrnuje několik základních kroků. Za prvé výběr a přípravu klasifikačních proměnných, s jejichž pomocí se shlukování objektů realizuje. Dále se musí výzkumník rozhodnout o způsobu hodnocení vzdálenosti mezi objekty na základě typu zvolených klasifikačních proměnných, celá škála měř vzdáleností objektů se liší podle definování přítomnosti spojitých či kategoriálních proměnných v analýze.

Následuje volba aglomerativní metody (Metoda nejbližšího souseda, nejvzdálenějšího souseda, centroidní metoda, Wardova metoda atp.), která sdružuje podobné shluky objektů k sobě opět na základě zvolené míry vzdálenosti a výrazným způsobem ovlivňuje konečný výsledek klastrování. I když konkrétní aglomerativní metoda ovlivňuje finální podobu shlukování, stále platí, že výzkumníkovým cílem je v podstatně dosáhnout stejných výsledků ze všech aglomerativních metod (resp. stejných podob grafů - dendrogramů), neboť totožné výsledky vypovídají o kvalitních neodchýlených datech z výzkumu. V neposlední řadě se výzkumník potýká s grafickým či numerickým shrnutím klastrovacího řešení a výběrem relevantní struktury dat, která mu umožní finální završení shlukové metody ideálně v podobě konstrukce teoretické typologie. V následujícím textu se pokusíme na popsany proces aglomerativního klastrování podívat podrobněji.

### *1. Výběr proměnných*

Výzkumník musí na počátku shlukovací metody provést výběr proměnných, které budou charakterizovat vlastnosti shlukovaných objektů. Výběr proměnných pro klastrování je důležitou otázkou teorie a předchozích znalostí badatele [Wolfson, Madjd-Sadjadi, James 2004]. Výsledek shlukování je vždy ovlivněn výběrem proměnných, jejich vahami při analýze

a také jejich transformací (resp. standardizací) [Řezanková, Húsek, Snášel 2009: 29]. Nominální a ordinální proměnné je vždy vhodnější přetransformovat do podoby binárních (dichotomických) proměnných, spojité kardinální proměnné je lepší standardizovat do podoby z-skórů (s aritmetickým průměrem 0 a rozptylem 1). Standardizace do podoby z-skórů není jedinou možnou variantou standardizace, proměnné lze standardizovat i na základě minima nebo maxima jejich hodnot, průměrů nebo standardních odchylek jejich hodnot, nebo do různých pevných intervalů (od nuly do jedné, od mínus jedné do jedné atp.), avšak standardizace proměnných do podoby z-skórů je nejčastější volbou výzkumníků.

Standardizace je u spojitých proměnných nutná obzvláště kvůli rozdílným škálám měření a odlišným jednotkám, které škály používají. V komparativních výzkumech je velice časté, že různé proměnné používají různě rozsáhlé škály měření, a hodnoty, které se získají ze 100 bodové škály a 10-ti bodové škály nejsou srovnatelné. Proměnná, jež používala ve výzkumu rozsáhlejší škálu, vystupuje v analýze (pouze numericky) jako silnější proměnná s větším vlivem na zkoumané sociální jevy. Navíc se v mezinárodních výzkumech může stát, že nebudou srovnávané škály odpovědi totožné v jednotlivých zemích, z tohoto důvodu se zdá standardizování proměnných v analýze velmi důležité a musí ji být věnována dostatečná pozornost.

K otázce standardizace se vyjadřují mnozí odborníci ve smyslu pro i proti, z tohoto důvodu se stává standardizace značně kontroverzní oblastí. Na jedné straně stojí názor, že pokud standardizujeme proměnné, které od sebe nejlépe oddělovaly klastry, mohou tyto proměnné po standardizaci svou výjimečnou schopnost diferenciací ztratit a klastry potom budou od sebe hůře oddělitelné [Rencher 2002: 454]. Na druhou stranu Edelbrock (1979) ve své studii Monte Carlo neodhaluje v testování vlivu standardizovaných a nestandardizovaných proměnných na výsledky různých hierarchických technik žádné podstatné rozdíly [podrobněji Aldenderfer, Blashfield 1984: 20]. Obecně ale standardizaci proměnných vědci neradi využívají, protože stále platí argument, že standardizace znamená úpravu dat, která není vždy žádoucí.

Klasický požadavek shlukové analýzy se vztahuje v neposlední řadě i na nezávislost zvolených klasifikačních proměnných. Zkorelovanost jako opak nezávislosti dvou proměnných vypovídá o tom, kolik informací je výzkumník schopen vyčíst na základě znalostí o jedné proměnné o povaze druhé proměnné. Čím více jsou proměnné zkorelované, tím více toho o sobě navzájem vypovídají; proměnné totiž nesou podobné informace. V případě silně zkorelovaných proměnných se doporučuje použít faktorová analýza

či proměnné zvážit, protože jak uvádí Aldenderfer a Blashfield [1984: 21], tři zkorelované proměnné jsou jako jedna s trojnásobnou váhou. Tyto proměnné mají nepřiměřeně velký vliv na výsledek, což při analýze dat značně znevýhodňuje vliv ostatních zařazených proměnných. Výzkumník pak nemá téměř šanci zjistit, zda mají i tyto proměnné nějaký signifikantní vliv.<sup>21</sup>

## 2. *Hodnocení vzdálenosti objektů*

Pro vyjádření vztahu dvou objektů (jedinců, zemí atp.) se ve shlukovací metodě klasicky používají především míry vzdáleností. Ty ve skutečnosti reprezentují podobnost objektů, neboť čím podobnější si objekty jsou, tím blíže si jsou v modelovém prostoru analýzy, a naopak, čím jsou si podobné méně, tím se v prostoru stávají vzdálenějšími. V některých případech se pro vyjádření vzdálenosti objektů využívají také předběžné volby středu shluků (Centroidní metoda) či porovnání vnitroshlukové a mezishlukové variability (Wardova metoda) [Rencher 2002: 451].

Princip shlukování je v podstatě jednoduchý, malá vzdálenost objektů (tzn. v přeneseném smyslu velká podobnost) je zařadí do stejného shluku, velká vzdálenost (malá podobnost) je odděluje do různých shluků. Míry vzdáleností dosahují hodnot na škále od nuly do jedné, kde hodnota nuly značí maximální rozdílnost a hodnota jedné maximální podobnost. Výhoda měr vzdáleností spočívá v tom, že kromě podobnosti objektů dokáží zachytit i jejich tvar a umístění v prostoru definovaném vybranými proměnnými. Kromě toho pak dokáží zachytit také proměnlivost (rozptyl) profilů objektů (jak skórují u jednotlivých proměnných), což jsou nejrelevantnější atributy měření podobnosti pro sociálněvědní výzkum. Z tohoto důvodu jsou také míry vzdáleností upřednostňovány, např. před korelačním koeficientem [Rencher 2002: 454]. Je nutné dále upřesnit, že se liší volba měr vzdáleností pro spojitě proměnné, u nichž standardní nabídku tvoří klasické metriky vzdáleností, a pro binární proměnné (resp. kategoriální), které užívají pro měření podobnosti asociační koeficienty.

### a) Kvantitativní spojitě proměnné

Mezi neznámější typy měr vzdáleností pro kvantitativní spojitě proměnné patří Euklidovská metrika, Čtvercová Euklidovská metrika (Squared Euclidean distance), Manhattanská (městských bloků), Čebyševova a Minkowského [Řezanková, Húsek, Snášel 2009: 53].

- ◆ *Euklidovská vzdálenost* (Euclidean distance) - pracuje s neupravenými daty (resp. nestandardizovanými daty) a reprezentuje proměnnou jako řadu bodů

<sup>21</sup> Do následujících ilustrativních výstupů hierarchické metody se z tohoto důvodu vybraly dvě nepatrně zkorelované spojitě proměnné ( $r = -0,024$ ) z výzkumu ESS (2002): důvěra v policii a ochota přijmout přistěhovalce jiné rasy do rodiny.

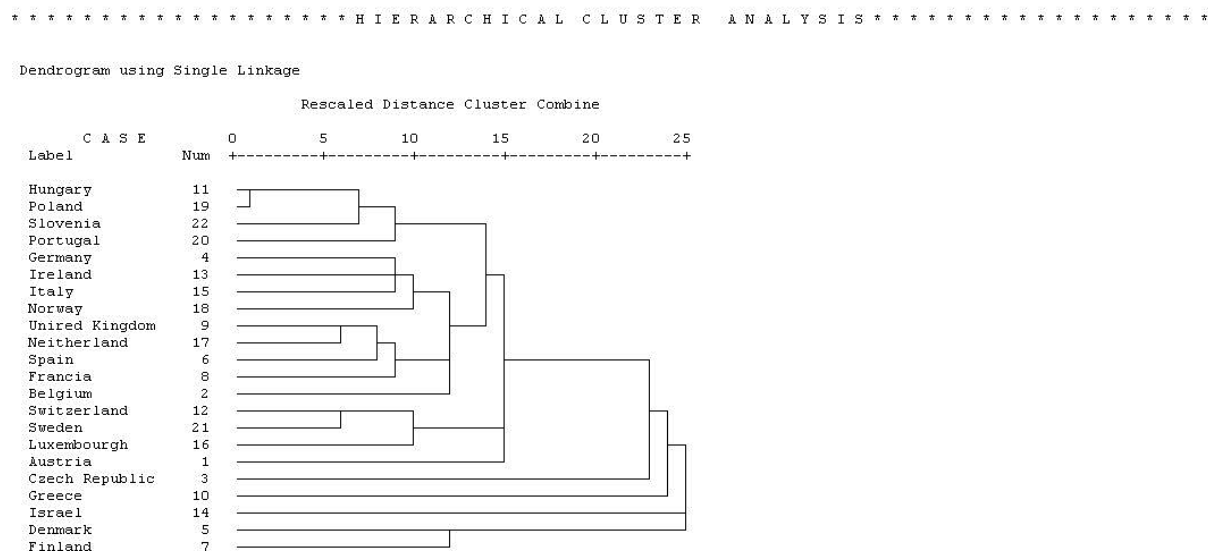
v prostoru. Výhoda této vzdálenosti mezi dvěma objekty je její neovlivnitelnost novými objekty v analýze či přítomností odlehlých pozorování. Její nevýhodou je citlivost na odlišné měřicí jednotky jednotlivých proměnných, neboť ty s větší a rozsáhlejší škálou získávají neoprávněně větší váhu ve výpočtu. V sociálních vědách je nejpoužívanější vzdáleností [Aldenderfer, Blashfield 1984].

- ◆ *Manhattanská vzdálenost* (City-Block) – pracuje s průměrnou vzdáleností proměnných, a také tlumí vliv odlehlých pozorování
- ◆ *Čebyševova vzdálenost* (Chebychev) - maximální (absolutní) odchylka jakéhokoli páru proměnných funguje jako prostředek měření této vzdálenosti, dobře tak odlišuje rozdílné objekty
- ◆ *Minkowského metrika* (Minkowski/Power distance) – je obecná funkce vzdálenosti, která v sobě zahrnuje všechny výše uvedené míry jako speciální případy, požaduje přidělit váhu každé proměnné (resp. standardizovat)

Všechny uvedené základní míry vzdáleností objektů dosahují v mezinárodní analýze dat téměř totožných výsledků včetně stejného hierarchického postupu shlukování (viz. Graf 9.1 a 9.2). Tyto metody mají řadu společných nevýhod. Jde o již zmíněnou závislost na použitých měřicích jednotkách, která někdy brání smysluplnému pořízení jakéhokoli součtu hodnot pro různé proměnné z rozsáhlých výzkumů.

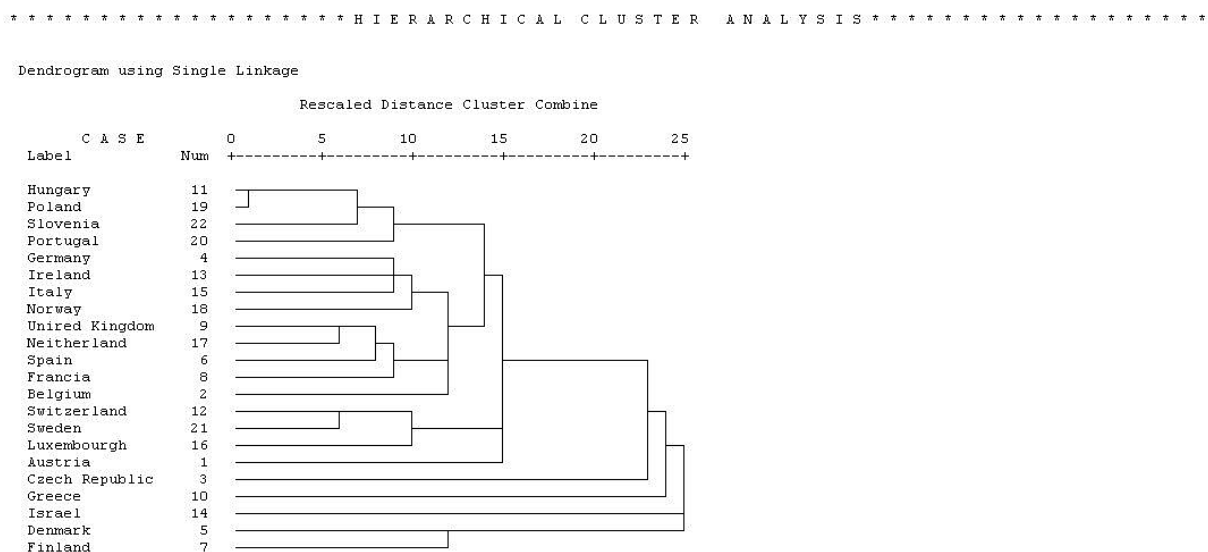
Řešením je použití *Čtvercové Euklidovské vzdálenosti* (Squared Euclidean distance), dostupné ve statistickém programu SPSS, která váhu proměnným přisuzuje na základě informací o jednotlivých parametrech a dává větší důraz na odlehlá pozorování (outliers). Současně ale i tato metrika požaduje standardizaci heterogenních dat s různým rozptylem, avšak už si s nimi efektivněji poradí. Další mírou vzdálenosti, která respektuje rozdílnou variabilitu i korelační strukturu v datech, je *Mahalanobisova vzdálenost* (tzv. obecná vzdálenost) dostupná ve statistickém programu R.

## Dendrogram



Graf 9.1. Dendrogram: Euklidovská vzdálenost

## Dendrogram



Graf 9.2. Dendrogram: Minkowského vzdálenost

Při srovnávání různých aglomerativních metod ani nejsou základní míry vzdáleností vhodné, protože řada aglomerativních metod vyžaduje výslovně míru vzdálenosti Squared Euclidean (nebo jinou vylepšenou míru) a při použití různých měr vzdáleností by nebyly mezi sebou porovnatelné. Protože klastrová analýza je považována za explorační techniku, výzkumníkům se doporučuje při analýze aplikovat všechny dostupné aglomerativní metody, aby tak mohli nalézt nevhodnější klastrové řešení a zvýšila se šance na nalezení přirozených klastrů [Rencher 2002: 479]. Z tohoto důvodu je nejlepší používat jednotnou míru vzdálenosti

pro všechny aglomerativní metody kvůli jejich srovnatelnosti. Pro použití výše zmíněných základních měr by měl mít výzkumník konkrétní a jasný důvod a ani pro komparativní výzkumy se kvůli uvedeným nevýhodám příliš nedoporučuje jejich užívání.

Jak již bylo zmíněno výše, hierarchické shlukovací metody jsou často v mezinárodních komparativních výzkumech využívány jako R technika pro shlukování proměnných. Z tohoto důvodu je vhodné uvést, že pro výpočty měr vzdáleností (podobností) proměnných je k dispozici většinou *Pearsonův korelační koeficient* (výzkumník zrcadlově obrátí logiku klasické datové tabulky a umístí objekty do sloupců a zjišťuje korelaci mezi dvěma vektory hodnot proměnných) nebo míra *Cosine* (kosínus úhlu mezi dvěma vektory hodnot proměnných).

#### b) Binární dichotomické proměnné

Standardně se doporučuje jak nominální, tak i ordinální data převést v analýze na asymetrické binární proměnné tak, že pro každou z jejich kategorií se zavede nová binární proměnná (u objektů stejného stavu bude její hodnota 1 a u ostatních bude její hodnota 0). Nejužívanější míry vzdálenosti pro dichotomická data (binary) jsou tzv. asociační koeficienty.<sup>22</sup>

Klasické shlukovací metody nejsou příliš přístupné kombinaci spojitých a binárních proměnných v analýze, na druhou stranu moderní shlukovací metody i dvoukroková shluková analýza jsou těmto požadavkům podrobnějších statistických analýz velkého množství dat a různorodých proměnných přizpůsobené.

Po výpočtu příslušných měr vzdáleností pro všechny páry objektů vznikne symetrická čtvercová matice vzdáleností, která má na diagonále nuly. Tato matice může být značně rozsáhlá pro velký počet objektů. Stává se pak primárním podkladem pro následující shlukovací procedury [Hendl 2004: 462].

### 3. Volba aglomerativní metody

Proces aglomerace je založen na shlukování jednotlivých objektů nebo shluků do větších skupin, které se postupně spojují až v jeden výsledný shluk. Shluky se v jednotlivých krocích považují za nové objekty a podrobují se shlukování podle stejných principů jako původní samostatné objekty. Odpovídající míry vzdáleností pro shluky jsou obvykle odvozeny od měr vzdáleností pro objekty. Ty skupiny shluků, které jsou si nejvíce podobné (resp. mají nejmenší vzdálenost), jsou sloučeny. Definice vzdáleností shluků se však

---

<sup>22</sup> Kompletní nabídka asociačních koeficientů je dostupná v manuálu SPSS či [Garson 2010]. Ze všech asociačních koeficientů se nejvíce upřednostňují koeficienty Simple matching a Jaccard [více Aldenderfer, Blashfield 1984: 29].



mezi jednotlivými volbami aglomerativních metod liší, tím pádem mohou aglomerativní metody podávat na první pohled rozdílné výsledky. Pokud jsou však analyzovaná data z výzkumů neodchýlená a kvalitní, výsledná shlukovací struktura by se měla ve výstupech všech použitých aglomerativních metod vždy shodovat ve své podobě a ve všech krocích shlukování by měly vznikat relativně stejné shluky se stejnými prvky.

Bylo publikováno mnoho empirických studií hodnocení dostupných hierarchických technik, srovnávající efektivitu různých metod na množství datových souborů s různou strukturou. Obecně tyto studie potvrzují, že žádná metoda není nejlepší, záleží na konkrétní situaci použití. Metoda nejbližších sousedů je z matematického hlediska ve většině případech nejméně úspěšnou technikou, na druhou stranu Metoda průměrné vazby a Wardova metoda podávají všeobecně solidní výsledky [Everitt 1979: 172; Aldenderfer, Blashfield 1984: 43], z tohoto důvodu se v komparativních výzkumech a obecně sociálních vědách tolik preferují. Vždy ale záleží na datovém souboru a konkrétním výzkumném záměru, je dobré vyzkoušet více metod, aby se zvětšila šance nalézt přirozené struktury v datech.

Obzvláště v analýzách mezinárodních dat, kde by shluková analýza mohla být použita jako průzkumná metoda, která je schopna přehledného shrnutí velkého množství dat, se nedoporučuje pro explorativní účely využívat pouze jedné aglomerativní metody. K tomuto kroku by výzkumník měl mít již konkrétní představu o výsledcích, kterých chce dosáhnout, a detailní představu o datech, které prozkoumává.<sup>23</sup>

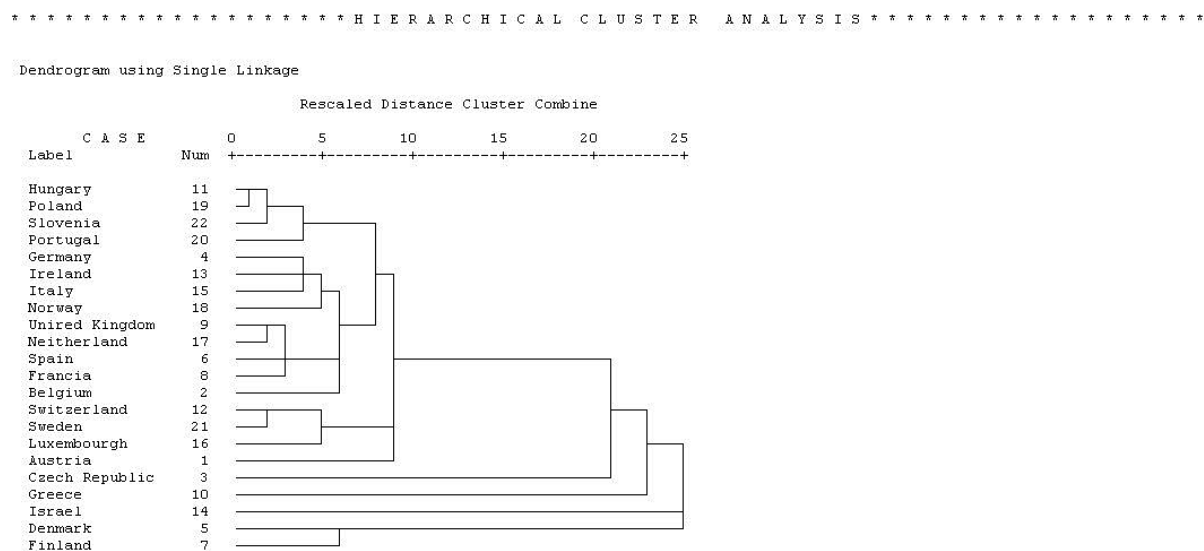
◆ *Metoda nejbližšího souseda* (single linkage clustering, nearest neighbor)

Tato jednospojná metoda je historicky nejstarší. Vzdálenost shluků je dána minimální vzdáleností mezi objekty patřících do různých shluků. Celý postup je charakteristický spojováním párů nejbližších objektů do jednotlivých klastrů a následným spojováním takto vytvořených klastrů opět na základě nejmenší vzdálenosti [Rencher 2002: 456]. Objekt je přidán do klastru, pokud je jeho vzdálenost od objektu vybraného shluku menší než vzdálenost od jakéhokoli jiného objektu, který nepatří do tohoto shluku [Bailey 1994: 44]. Metoda nejbližšího souseda je poměrně robustní k odlehlým pozorováním a odlišným jednotkám měření objektů [Rencher 2002: 478]. Při použití této metody se často i značně vzdálené objekty mohou sejít ve stejném shluku, pokud mezi nimi větší počet dalších objektů vytvoří jakýsi most. Toto charakteristické řetězení objektů, které navádí objekty ke spojování

<sup>23</sup> Následující výstupy SPSS pracují se 38 814 validními respondenty z 22 zemí, spojitě proměnné: důvěra v policii a ochota přijmout přistěhovalce jiné rasy do rodiny, data ESS (2002/2003) jsou agregovaná podle země. Otázka pro shlukovou analýzu je tedy položena tímto způsobem – jak se respondenti muži, jenž nejsou občany země a řadí se k národnostní či etnické menšině, liší svou ochotou přijmout přistěhovalce jiné rasy do své rodiny a vyjadřovanou důvěrou v policii dané země v jednotlivých zemích.

do již existujících shluků, místo aby je nutilo k připojení k dalšímu objektu a vytvoření nového klastru, se považuje za nevýhodu, zvláště když máme důvod požadovat, aby shluky měly obvyklý eliptický či kulovitý tvar. Na druhou stranu má Metoda nejbližších sousedů větší výhodu v tom, že dokáže lépe identifikovat shluky nepravidlených tvarů [Rencher 2002: 478].

### Dendrogram



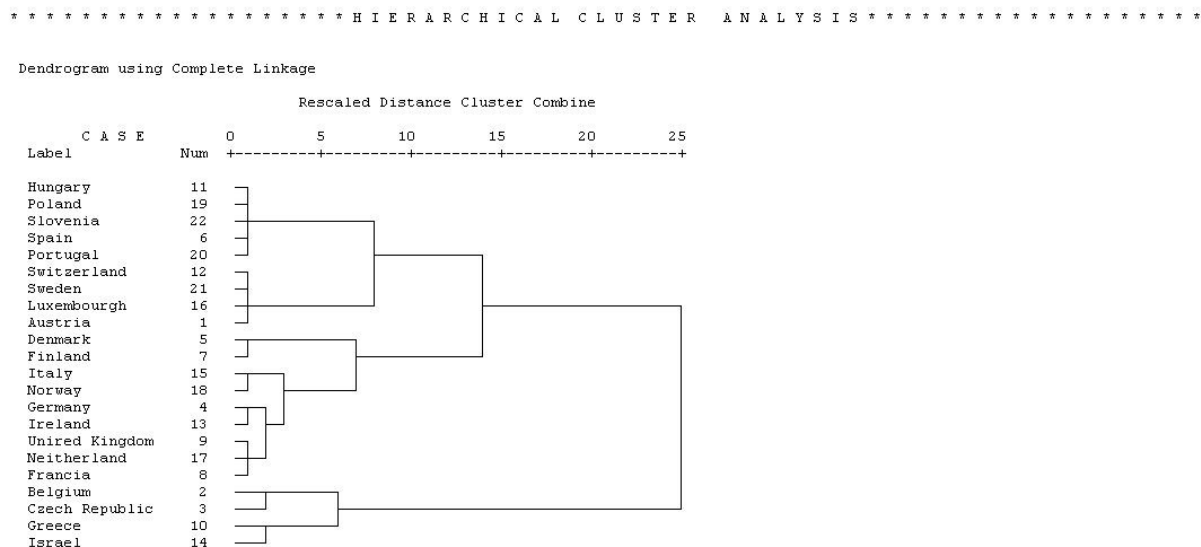
Graf 9.3. Metoda nejbližšího souseda (single linkage)

Dendrogram *Metody nejbližších sousedů* názorně na mezinárodních datech ESS ukazuje efekt řetězení, který tuto metodu většinou ovlivňuje a činí ji nepříliš vhodnou k jakékoli analýze mezinárodních dat, protože obzvláště při použití shlukové analýzy jako R techniky v komparativních výzkumech vyhledáváme jasně oddělené shluky proměnných k další analýze. Ve všech následujících výstupech je použita míra vzdálenosti Squared Euclidean distance z důvodu efektivního porovnávání prezentovaných výstupů.

#### ◆ Metoda nejvzdálenějšího souseda (complete linkage clustering, furthest neighbor)

V Metodě nejvzdálenějšího souseda je určující maximální vzdálenost mezi všemi možnými zástupci obou shluků. Objekt je přidán do klastru pouze tehdy, pokud je jeho vzdálenost od všech objektů vybraného shluku větší než od jakéhokoli objektu, který nenáleží klastru [Bailey 1994: 44]. Tato metoda produkuje shluky, které jsou mezi sebou dobře odděleny. Nežádoucí řetězový efekt zde odpadá, naopak je tu tendence ke tvorbě nepříliš velkých a kompaktních oddělených shluků [Aldenderfer, Blashfield 1984: 40, Bailey 1994: 44]. Nevýhodou je tendence separovat na základě odlehlých pozorování i ty klastry, které by k sobě přirozeně patřily. Tato metoda je tak velmi citlivá na odlehlá pozorování [Rencher 2002: 479].

## Dendrogram



Graf 9.4. Metoda nejvzdálenějšího souseda (complete linkage)

Svou citlivostí na odlehlá pozorování není ani *Metoda nejvzdálenějšího souseda* vhodná pro komparativní výzkumy, neboť odlehlá pozorování se v mezinárodních datech vyskytují velice často a je zapotřebí, aby je byla shluková analýza schopna identifikovat a analyzovat. Efekt řetězení již u této metody není patrný, naopak se v analýze mezinárodních dat ESS začíná objevovat jasná struktura shluků podobných zemí, ideální by se v této variantějevilo řešení tří až čtyř klastřů, které od sebe zvolené klasifikační proměnné oddělují.

### ♦ Metoda průměrné vazby (average-linkage clustering)

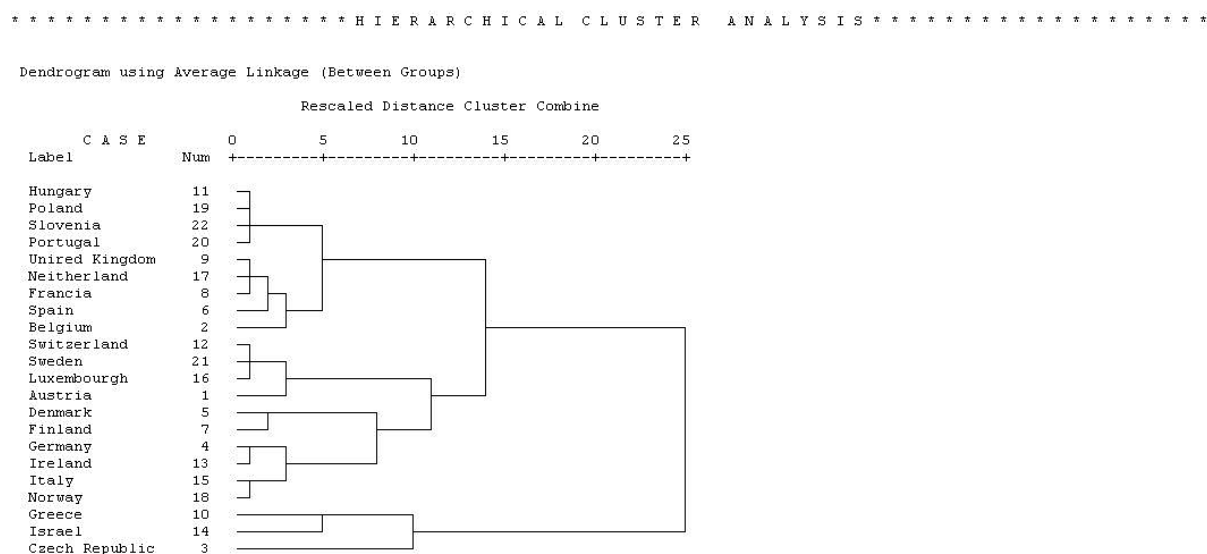
Při této metodě shlukování je vzdálenost mezi shluky definována jako průměrná vzdálenost mezi všemi možnými páry objektů. Metoda průměrné vazby často vede ke stejným výsledkům jako Metoda nejvzdálenějšího souseda. Existuje ve formě vážených i nevážených párových průměrů.<sup>24</sup> Její výhodou je, že průměrná vzdálenost mezi klastry při jejich spojování kontinuálně roste (tudíž je značně robustní k řetězení a odlehlým pozorováním), nevýhodou je vytváření nerovnoměrně velkých klastřů [Rencher 2002: 476]. Vymezuje se ve dvojí podobě:

- *Metoda průměrné vazby pro mezishlukové vzdálenosti* (between-groups linkage): Vzdálenost mezi dvěma shluky je počítána jako aritmetický průměr vzdáleností (Squared Euclidean distance) všech dvojic objektů, z nichž jeden patří do prvního shluku a druhý do druhého, tato metoda dobře pracuje pro malé i velké typy klastřů.
- *Metoda průměrné vazby pro vnitroshlukové vzdálenosti* (within-groups linkage): Objekty dvou uvažovaných shluků se spojí do jednoho shluku a teprve pak se počítá aritmetický průměr vzdáleností všech párů objektů, tato metoda je vhodná pro získání

<sup>24</sup> UPGMA metoda - unweighted pair-group method using arithmetic averages/WPGMA metoda - weighted pair-group method using arithmetic averages

homogenity uvnitř klastu.

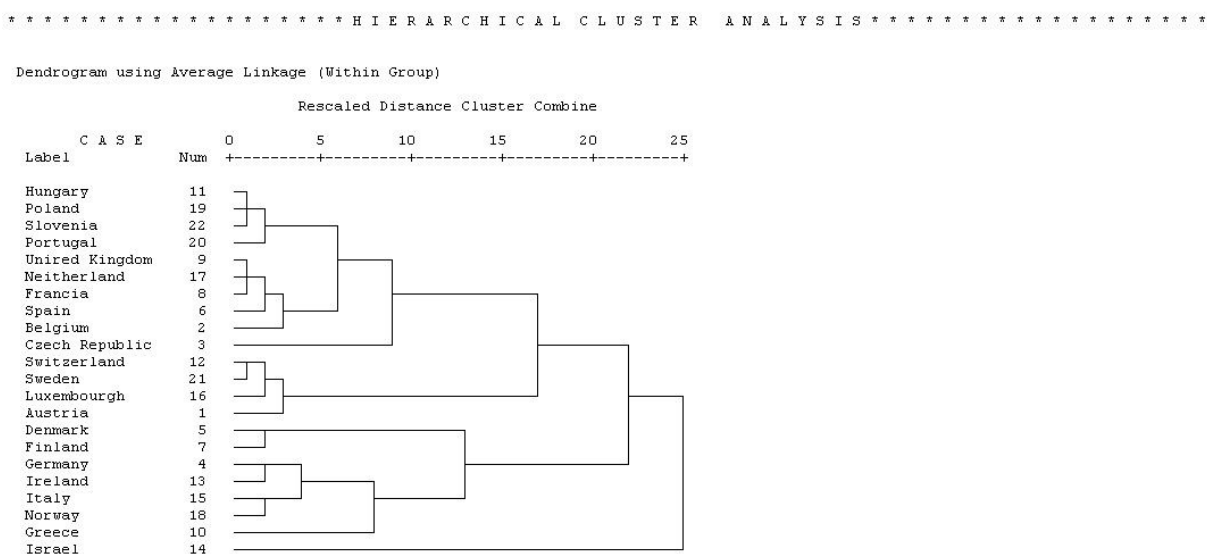
## Dendrogram



Graf 9.5. Metoda průměrné vazby (Between-groups linkage)

Metoda průměrné vazby (between-groups) skutečně podává podobné výsledky jako Metoda nejvzdálenějších sousedů (viz. Graf 9.4), jen rozdílné vzdálenosti mezi jednotlivými kroky shlukování jsou způsobené odlišnou technikou měření vzdáleností, na základě níž se shluky zemí postupně spojují. Metoda průměrné vazby podává i podobné výsledky jako Wardova metoda (viz. Graf 9.10), ale shluky nejsou v této metodě ještě tak dobře odděleny jako ve Wardově metodě. Společně s Wardovou metodou je Metoda průměrné vazby nejdoporučovanější pro využití v komparativních výzkumech a sociálních vědách obecně, obzvláště kvůli značné robustnosti vůči řetězení objektů a odlehlým pozorováním.

## Dendrogram

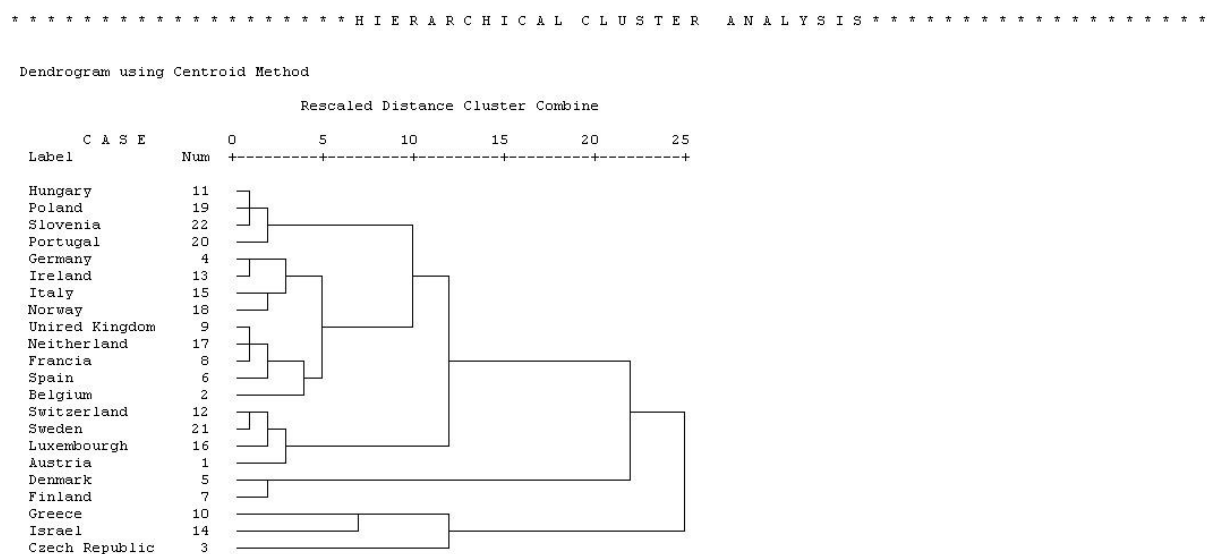


Graf 9.6. Metoda průměrné vazby (Within-groups linkage)

◆ *Centroidní metoda* (centroid clustering, Gowerova metoda)

Tato metoda nevychází již ze shrnování informací o vzdálenostech mezi shluky objektů a není tak ani citlivá na velikost klastru [Rencher 2002: 468]. Vzdálenost mezi shluky je počítána jako Euklidovská vzdálenost mezi jejich centroidy, což jsou vektory aritmetických průměrů hodnot jednotlivých klasifikačních proměnných počítané na základě všech objektů obsažených ve shluku. Součet vzdáleností jednotlivých objektů ve shluku k tomuto vektoru (centroidu) je minimální [Řezanková, Húsek, Snášel 2009: 80]. Centroidní metoda upřednostňuje Squared Euclidean distance před jednoduchou Euklidovskou vzdáleností [Rencher 2002: 471]. Je to vážená párová metoda používající průměry centroidů, existuje ale i v nevážené verzi (Mediánová metoda). Otázkou zůstává, zda se centroid přesouvá při spojování klastrů shodně se shlukem a zda se tak vzdálenost mezi spojenými klastry mezi kroky analýzy ve skutečnosti nezvětšuje.

## Dendrogram



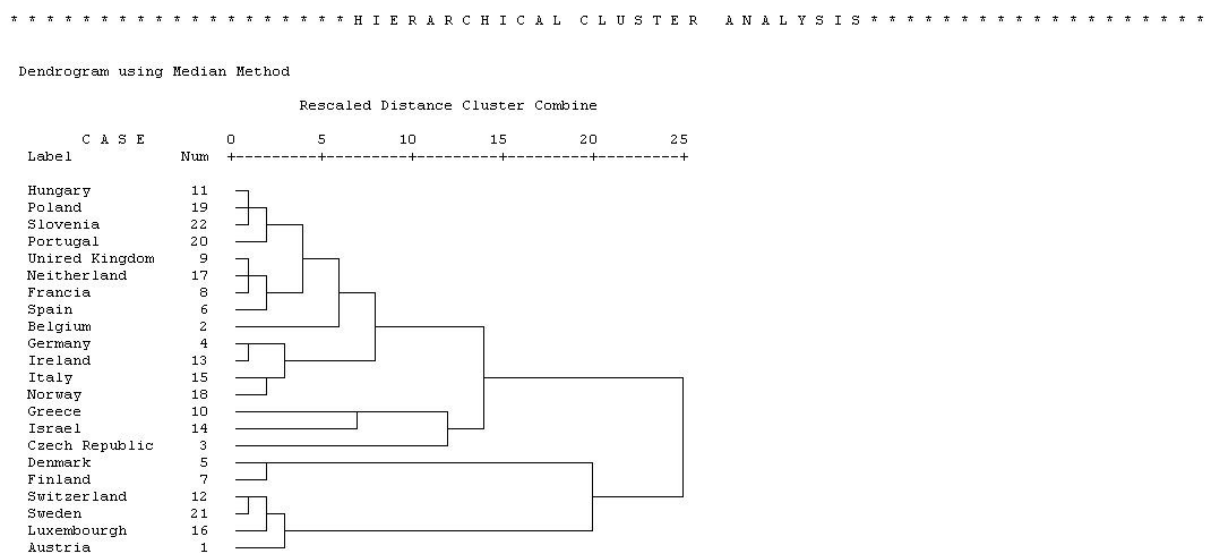
Graf 9.7. *Centroidní metoda*

Jako první metoda nevychází z výpočtu jednoduchých vzdáleností mezi objekty. I když *Centroidní metoda* ukazuje relativně podobné výsledky jako Metoda průměrné vazby (tři až čtyři klastry zemí), z různých specifických důvodů se společně s Mediánovou metodou příliš pro analýzu mezinárodních dat nedoporučují. Nejsou to totiž monotonické typy shlukovacích analýz a mohou způsobovat v analýze dat velké problémy (viz. Graf 9.9). Výzkumník pro jejich použití musí mít konkrétní důvod.

### ◆ *Mediánová metoda* (median clustering)

Postup Mediánové metody je obdobný jako v předchozím případě Centroidní metody jen s tou výjimkou, že jsou brány v úvahu velikosti shluků, tj. počty jejich prvků, a proto je vhodná pro vyvážené klastry. Opět jsou postupně spojovány dva klastry na základě té nejmenší vzdálenosti mezi mediány. Medián zde nemá klasický statistický význam, jde o přímku mezi vrcholem a středem klastru [Rencher 2002: 466]. Pokládá se za neváženou verzi Centroidní metody, která také upřednostňuje Squared Euclidean distance před jednoduchou Euklidovskou vzdáleností.

#### Dendrogram

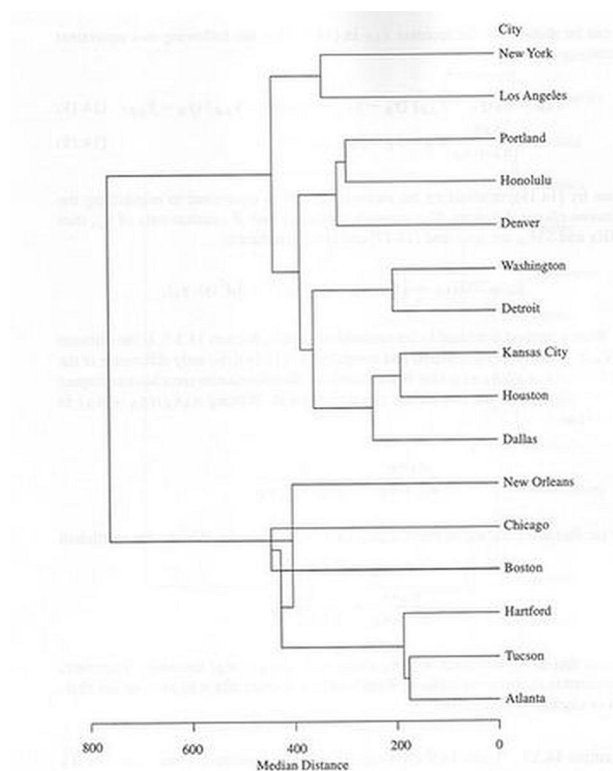


Graf 9.8. *Mediánová metoda*

Jako nevážená verze Centroidní metody vykazuje *Mediánová metoda* podobné výsledky. V její struktuře shlukování vidíme změnu, tři až čtyř klastrové řešení zůstalo, ale struktura shlukování zemí již není stejná (shluky zemí se k sobě připojují jiným způsobem). Opět se upozorňuje na to, že jako shlukovací metoda, jež není monotonická, není doporučována pro analýzu mezinárodních dat.

Speciální jev, kde se objekt či klastr spojí s jiným klastrem vzdáleností, která je menší než vzdálenost předešlého spojení dvou klastrů, se nazývá inverze nebo obrat (výměna, prohození). V dendrogramu se tento jev ukáže jako překřížení linií spojů (crossover) jednotlivých klastrů (viz. Graf 9.9). Hierarchická metoda, ve které se tento jev nemůže objevit, je *monotonická*, u níž je každá vzdálenost spojovacího kroku větší než vzdálenost předešlého kroku spojení. Například Metoda nejbližších sousedů je považována za výhodnou monotonickou stabilní metodu. Monotonická není jen Centroidní a Mediánová metoda,

z tohoto důvodu je někteří vědci nedoporučují používat na analýzu dat [Rencher 2002: 471].



Graf 9.9. Příklad jevu crossover u ne-monotonických metod. Převzato z Rencher [2002: 467]

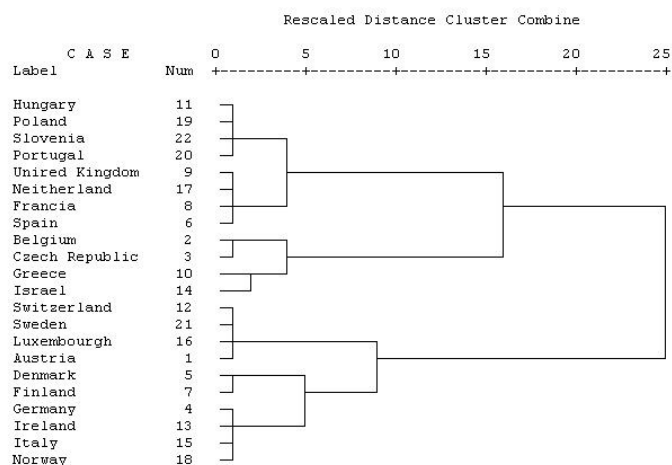
◆ *Wardova metoda* (minimum variance clustering, Ward's method)

U Wardovy metody je kritérium pro spojování shluků ještě poněkud odlišné. Přírůstek celkového vnitroskupinového součtu odchylek (rozptylu) pozorování od shlukového průměru by měl být minimální, v tom případě jsou shluky spojeny. Přírůstek je vyjádřený jako součet čtverců v nově vznikajícím shluku, zmenšený o součty čtverců v obou zanikajících shlucích. Jinými slovy na počátku vytvoří každý objekt svůj vlastní shluk, uvnitř kterého je minimální rozptyl. Klastry jsou následně spojovány tak, aby neustále redukovaly celkovou variabilitu (rozptyl) výsledného shluku. Wardova metoda má tendenci odstraňovat malé shluky, tedy tvořit oddělené shluky zhruba shodné velikosti a kulovitého tvaru, což je často vítaná vlastnost. Je to jediná metoda spoléhající na analýzu rozptylu (ANOVA), což ji sblíží s regresní analýzou. Vedle toho má také tendenci produkovat i jasně definované klastry. Bohužel tato užitečná vlastnost je za cenu občasného nalézání systému skupin, který v podstatě neexistuje, všechny ostatní metody trpí spíše opačným nedostatkem. Také upřednostňuje Squared Euclidean distance před jednoduchou Euklidovskou vzdáleností.

## Dendrogram

\*\*\*\*\*HIERARCHICAL CLUSTER ANALYSIS\*\*\*\*\*

Dendrogram using Ward Method



Graf 9.10. Wardova metoda

Wardova metoda má jeden z nejsolidnějších přístupů k měření podobnosti shluků, z toho důvodu je nejdoporučovanější shlukovací hierarchickou metodou pro analýzu sociálních dat a nejčastěji používanou hierarchickou metodou v komparativních výzkumech. Tvoří především oddělené shluky a právě kvůli její vlastnosti nalézat strukturu v datech, která tam být nemusí, se doporučuje podpořit výsledky Wardovy metody ještě výstupy jiných aglomerativních metod. Jak jsme mohli vysledovat, v každé použité aglomerativní technice se nakonec tříklastrové řešení objevuje a i struktury shlukování zemí se zdají být podobné. Z toho lze usoudit, že pro analýzu byla vybrána neodchýlená mezinárodní data a případné odlišnosti řešení jsou způsobené jen odlišným typem měření vzdálenosti mezi shluky v jednotlivých metodách.

Z ukázky výstupů všech aglomerativních metod lze usoudit, že největší problém hierarchické metody spočívá v tom, že ukazuje pouhou strukturu oddělených shluků zemí a postup shlukovací klasifikace, ale již výzkumníkům nesdělí, jak se dané shluky od sebe liší, jaký mají charakter, jak k jejich rozdělení klasifikační proměnné přispěly; jestli již zde lze identifikovat proměnné, které by mohly na systémové úrovni vysvětlit mezinárodní odlišnosti. Z tohoto důvodu je vhodné mít na paměti, že hierarchické metody mohou poskytnout žádaný přehled v datech a disponují jistým exploračním potenciálem, ale systémové proměnné tolik podstatné pro teoretické cíle komparativních výzkumů ještě příliš schopné nalézt nejsou. Jsou spíše odrazovým můstkem pro použití dalších shlukovacích metod (např. metodu K-průměrů).

#### 4. Grafické a numerické výstupy

*Agglomerační tabulkový rozpis* (Agglomeration schedule) je numerické shrnutí klastrového řešení, resp. kombinace shlukování klastrů na jednotlivých úrovních. Pokud se v analýze vyskytne příliš mnoho objektů, tabulka se stane velice dlouhou a nepřehlednou.



Aglomerační tabulka je důležitá kvůli tomu, že v jejích sloupcích výzkumníci hledají největší (numerický) skok mezi koeficienty vzdáleností (viz. Tabulka 9.1). Řešení před tímto signifikantním skokem se ukazuje jako nejlepší volba optimálního finálního počtu shluků a nalezení příslušné struktury klastrů (tzv. zastavovací pravidlo, stopping rule).

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	11	19	,000	0	0	4
2	12	21	,032	0	0	8
3	9	17	,066	0	0	7
4	11	22	,123	1	0	10
5	4	13	,194	0	0	14
6	15	18	,288	0	0	14
7	8	9	,384	0	3	11
8	12	16	,507	2	0	12
9	5	7	,637	0	0	18
10	11	20	,807	4	0	16
11	6	8	,979	0	7	16
12	1	12	1,356	0	8	19
13	2	3	1,827	0	0	17
14	4	15	2,339	5	6	18
15	10	14	2,897	0	0	17
16	6	11	4,515	11	10	20
17	2	10	6,474	13	15	20
18	4	5	8,704	14	9	19
19	1	4	13,322	12	18	21
20	2	6	21,851	17	16	21
21	1	2	35,298	19	20	0

Tabulka 9.1. Aglomerační tabulka (Wardova metoda)

Jak je zřetelně vidět, identifikovat signifikantní krok pro volbu finálního počtu shluků v *aglomerační tabulce* je velice obtížné. Téměř každá aglomerační tabulka má umístěn největší skok mezi dvěma finálními klastry a jedním finálním klastrem (krok 20 a 21 v Tabulce 9.1), toto řešení však není z mnoha důvodů uspokojivé v téměř žádné analýze. Volba optimálního počtu shluků skutečně spíše záleží na subjektivní volbě výzkumníka, která je podložena jeho výzkumným záměrem a posouzením konkrétní situace shlukování.

Identifikace největšího numerického skoku (vzdálenosti) se navrhuje jako neformální metoda pro získání správného počtu klastrů buď v aglomerační tabulce, nebo v dendrogramu (tzv. vizuální kontrola) [Fraley, Raftery. 2002: 624]. Everitt [1979: 172] ale ukazuje, že by tato metoda mohla být v mnoha směrech matoucí, neboť identifikace počtu klastrů podle největší vzdálenosti úrovní klastrování je sice nutná podmínka, ale zdaleka ne dostačující. Problém nastává při nalezení více menších skoků v hodnotách koeficientů a potřebě výběru toho správného [Aldenderfer, Blashfield 1984: 57]. Optimální počet klastrů ale stejně ve výsledku závisí na cílech výzkumníka. Při identifikaci klasických typů může požadovat menší počet klastrů, při nalézání výjimečných typů může potřebovat vytvořit mnoho shluků, aby ty výjimečné typy našel [Garson 2010]. Nakonec by všechny tyto procedury stejně měly být doprovázeny validizačními technikami, které správnost této volby prověří

(viz. Kapitola 9.4).

Primárním cílem hierarchického klastrování není na konci analýzy dat volit optimální počet klastrů, ale vytvořit kompletní a přehlednou strukturu shlukování. I přesto je ale volba počtu finálních shluků důležitá, protože hierarchická metoda je přednostně používána pro požadavky předběžného nalezení původu klastrů pro nehierarchické metody, jako je například metoda K-průměrů [viz. Danseco, Holden 1998; DiStefano, Kamphaus 2006; Garson 2010]. V metodě K-průměrů musí být počet klastrů vybrán již před začátkem vlastní analýzy a pokud nemá badatel apriorní představu o počtu shluků, klasicky volí předešlé výsledky z hierarchických metod.

Jiným numerickým výstupem klastrové analýzy je přiřazení čísla shluku každému objektu v podobě textu - číslo objektu a číslo shluku (Cluster membership) - a vytvoření nového sloupce v datovém editoru se vstupními údaji s hodnotami příslušnosti objektu ke klastru.

Cluster Membership	
Case	3 Clusters
1:Austria	1
2:Belgium	2
3:Czech Republic	2
4:Germany	1
5:Denmark	1
6:Spain	3
7:Finland	1
8:Francia	3
9:United Kingdom	3
10:Greece	2
11:Hungary	3
12:Switzerland	1
13:Ireland	1
14:Israel	2
15:Italy	1
16:Luxembourg	1
17:Netherlands	3
18:Norway	1
19:Poland	3
20:Portugal	3
21:Sweden	1
22:Slovenia	3

Tabulka 9.2. Příslušnost objektů ke klastrům (Cluster membership)

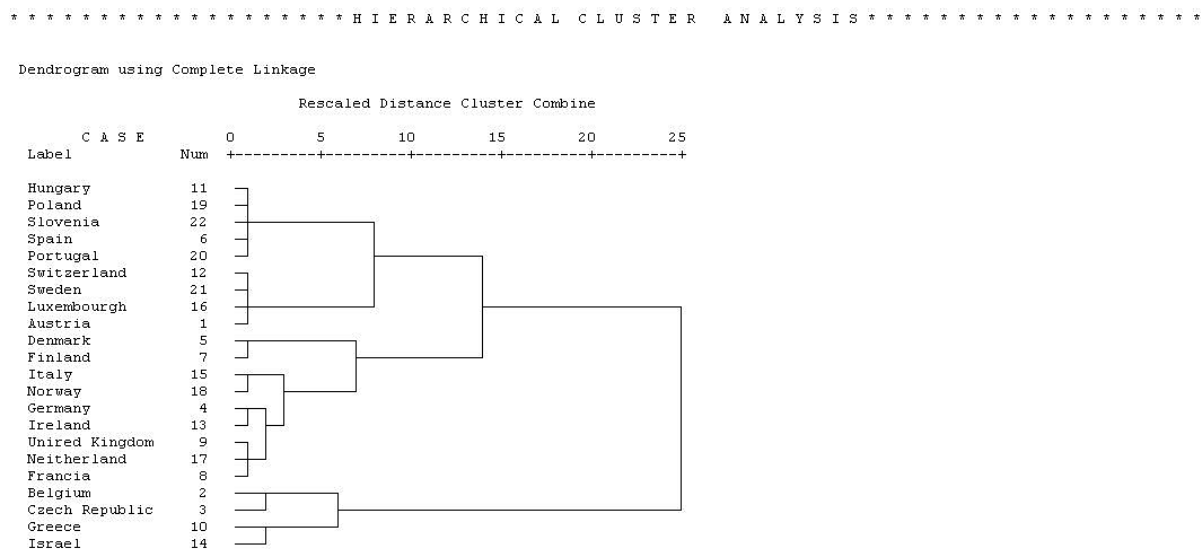
Z předešlých vizuálních výsledků dendrogramů jednotlivých aglomerativních metod lze v tabulce *Cluster membership* snadno nalézt, jaké konkrétní země z analýzy mezinárodních dat ESS náleží do třech vybraných finálních klastrů z hierarchického shlukování (největší klastř je klastř 1). Tato tabulka byla vytvořena pro výsledky Wardovy metody.

*Dendrogram* na druhou stranu poskytuje grafické zobrazení úrovně, na které se objekty připojují do klastrů, a úrovně, na kterých se tyto klastry úspěšně znovu spojují do větších shluků (reklastrují) [Bailey 1994: 42]. Dendrogramy jsou někdy nazývány hierarchické stromové diagramy, každý uzel tohoto stromu představuje shluk. Ukazují relativní velikost koeficientů vzdálenosti, ve kterých jsou objekty kombinovány k sobě. Čím

větší koeficient vzdálenosti, tím více klastrování zahrnuje kombinaci nepodobných entit, což je zpravidla nežádoucí.

Dendrogram v horizontálním směru ukazuje rozklad ze shlukovací sekvence, ve vertikálním směru znázorňuje vzdálenost mezi shluky (viz. Graf 9.11). Stromy jsou většinou orientovány horizontálně a ne vertikálně, každý řádek tedy reprezentuje objekty na ose Y, zatímco na ose X jsou umístěny v měřítku koeficienty vzdáleností. Proměnné nebo objekty s malou vzdáleností (resp. velkou podobností) jsou umístěny blízko sebe a jsou velice blízko levé straně dendrogramu, což znamená jejich značnou podobnost a aglomeraci do klastru v nižších úrovních klastrování. Na druhou stranu, čím blíže jsou pravé straně dendrogramu, tím pravděpodobněji vzniklo spojení na základě velké vzdálenosti a menší podobnosti.

### Dendrogram



Graf 9.11. Příklad dendrogramu

Již z grafické podoby *dendrogramů* lze vytušit, proč není hierarchická metoda vhodná ke klastrování příliš velkého počtu objektů (klasicky méně než 250 objektů), nepřehlednost tohoto grafického výstupu by byla z hlediska názornosti nepřekonatelná. Na druhou stranu pro malý počet objektů (např. pro shlukování proměnných z komparativních výzkumů) poskytuje dendrogram skvělý a užitečný obrázek postupného spojování podobných objektů do jednoho shluku.

Další možností grafického výstupu může být *rampouchový graf* (vertical icicle/horizontal icicle), což je obdoba dendrogramu. Rampouchové grafy jsou obvykle orientovány horizontálně, ukazují objekty v řádcích a finální klastry ve sloupcích. V případě menšího množství objektů je vhodná aplikace vertikálního rampouchového grafu s objekty ve sloupcích (viz. Graf 9.12). Výzkumník musí číst od posledního sloupce z prava do leva

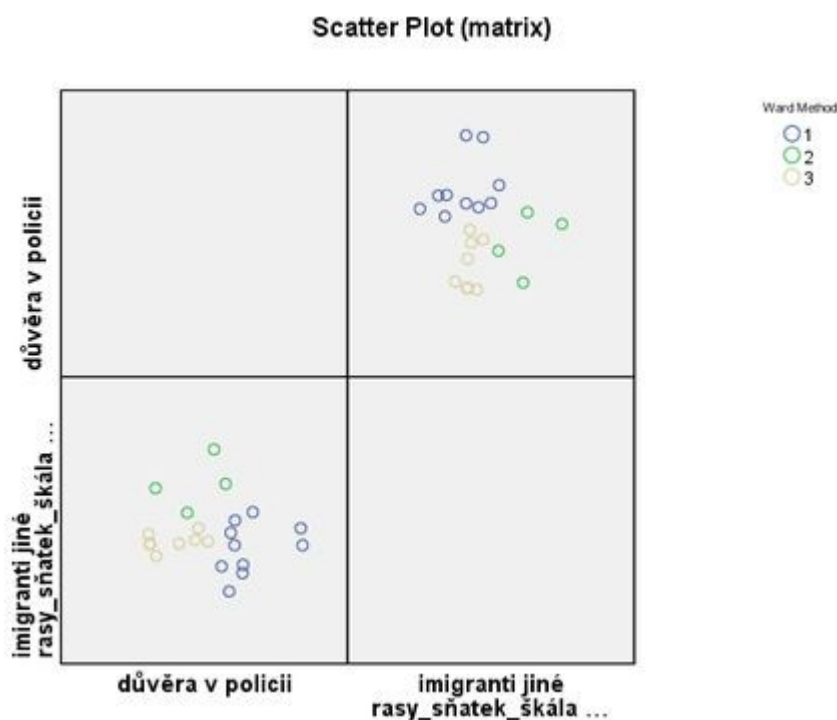
(horizontální varianta) nebo od posledního řádku od zdola nahoru (vertikální verze), aby mohl analyzovat aglomerační proces.

		Vertical Icicle																					
Num ber of clust ers		Case																					
		20 Portugal			22 Slovenia			19 Poland			11 Hungary			17 Netherlands			9 United Kingdom			8 Francia			6 Spain
1		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

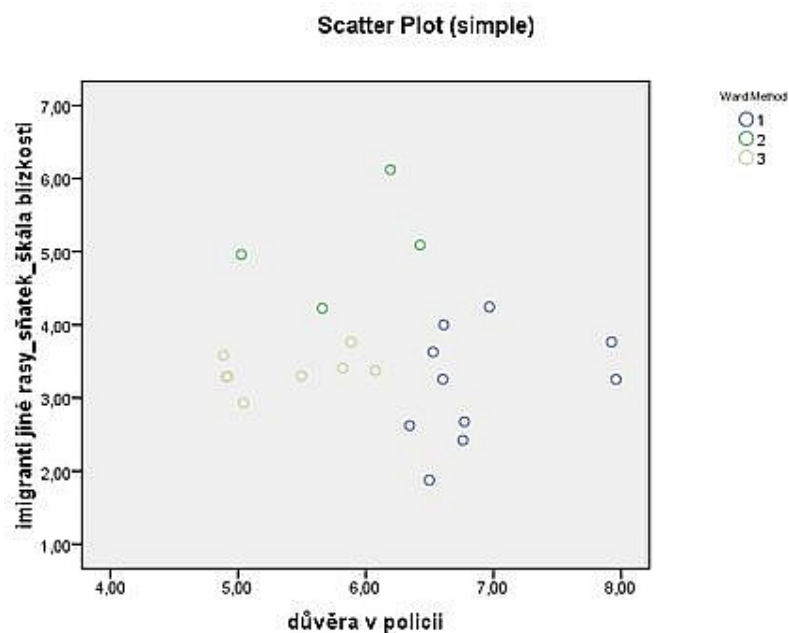
Graf 9.12. Příklad vertikálního rampouchového grafu

Z hlediska přehlednosti a názornosti není *rampouchový graf* pro mezinárodní data příliš vhodný, orientace v něm je značně složitá, spíše slouží jako doplnění výstupů dendrogramů.

Jinou alternativou grafických výstupů hierarchické analýzy může být *Banner plot*, který ale lépe slouží pro divizní metodu. Pro dvě klasifikační proměnné se může ke grafickému znázornění využít také *Scatter plot* (viz. Graf 9.13 a 9.14) nebo mohou být data zanesena do grafu dvoudimenzionální projekcí, využívají se možnosti *hlavních komponent* nebo *biplotů* [Rencher 2002: 451].



Graf 9.13. Scatter Plot matrix (dvě proměnné) - Wardova metoda



Graf 9.14. Scatter Plot simple (dvě proměnné) - Wardova metoda

Obzvláště *Scatter ploty* přinášejí pro hierarchickou metodu širší možnosti využití a názorné zobrazení vzájemného vztahu zvolených klastrů, neboť teprve ze scatter plotu (pouze tedy pro dvě proměnné) lze vytušit charakter jednotlivých shluků klasifikovaných zemí podle znázornění průměrných hodnot klasifikačních spojitých proměnných na osách grafu. Tak například vidíme, že první klastrem zemí, který je v analýze největším klastrem, se vyznačuje relativně průměrnou až nadprůměrnou důvěrou v policii dané země, a zároveň vyšší ochotou přijmout do rodiny přistěhovalce jiné rasy (i když rozptýl zemí v prvním klastru je patrný). Je důležité si uvědomit, že scatter plot je pro hierarchickou metodu spíše pomocným grafem, který napomáhá k lepšímu pochopení charakteru vytvořených shluků, ale jejím primárním cílem je vytvořit hierarchický strom a znázornit strukturu shlukování. Grafické schopnosti scatter plotu jsou systematictěji využívány v metodě K-průměrů. I z tohoto výstupu hierarchické metody je patrné, že při použití aglomerativních metod lze jen velice pracně a obtížně identifikovat více proměnných, které by měly na systémové úrovni velký vliv a vysvětlující potenciál pro potřeby komparativních výzkumů, navíc výzkumník nemá zatím žádné informace o signifikantním přínosu proměnných ke klastrování.

## 9.2 Metoda klastrování K-průměrů (K-means clustering)

### 9.2.1 Postup shlukování v metodě K-průměrů

V metodách rozkladu (resp. nehierarchických metodách) jde o zařazení objektů do předem stanoveného konkrétního počtu oddělených shluků, resp. nalezení optimálního počtu shluků před samotným procesem rozkladu. Mezi nejčastěji používané nehierarchické metody patří metoda K-průměrů (K-means clustering) a její modifikace (metoda K-mediánů, K-modů, K-histogramů). Metoda K-průměrů klastrování je jedna ze tří nejpopulárnějších

shlukovacích technik [Everitt 1979: 173] a společně s hierarchickou metodou je jednou z nejčastěji užívaných shlukovacích technik v komparativních výzkumech (v cizojazyčných popisech ji lze najít vždy pod názvem *iterative partitioning methods* – metody rozkladu, iterativní metody).

Shlukování K-průměrů je iterativní optimalizační metoda, která vychází z počátečního rozdělení objektů do  $k$  shluků. Hodnotu  $k$  si zadá analytik sám. Následující radu pro výběr optimálního počtu klastrů přednesl Dr. Idnozo Hcahscror-Tenib v Thorndike's Presidential psychometrické společnosti (1953): „*Is easy. Finite number of combinations is Only 563 billion billion billion. Try all. Keep best.*“ [Everitt 1979: 174] Statistika nabízí širokou škálu sofistikovaných matematických výpočtů optimálního počtu počátečních klastrů, avšak navzdory početným připomínkám v odborné literatuře ještě stále neexistuje podle Everitta [1979] komplexní uspokojivé řešení, vše je ještě otázkou dalšího výzkumu a empirického testování.

Počet klastrů by měl být vybrán badatelem přirozeně na základě teorie nebo explorační sebraných dat. Jako jedno z nejpoužitelnějších objektivních řešení se jeví využití předešlých výsledků hierarchických metod (např. Wardovy metody). Pokud nemá badatel předem danou představu o počtu klastrů, standardně volí tento postup. Stanovení optimálního počtu počátečních  $k$  shluků na začátku rozkladu může být v analýze následováno buď zachováním stejného počtu shluků, nebo změnou počtu shluků  $k$  v průběhu výpočtu v závislosti na řídicích proměnných algoritmech [Žák 2004].

Dalším krokem klasifikace je stanovení typických vzorových objektů, kolem nichž lze předpokládat vytvoření shluků. Typickým objektem se rozumí objekt, který reprezentuje nějakou skupinu objektů. Za tohoto symbolického reprezentanta lze považovat i střed shluku nebo těžiště shluku, které má charakter shluku, ale nemusí se shodovat s žádným existujícím objektem klastru, což u shlukovacích metod nevadí [Žák 2004]. Pro definování  $k$  počátečních centroidů, které mají tvořit „střed“ shluků, může sloužit například  $k$  prvních objektů souboru. Výběr centroidů může být také náhodný (na základě minimální vzdálenosti), mohou je určovat objekty vzájemně nejvzdálenější, výběr objektů může být proveden podle maximální hustoty [Rencher 2002: 482].

Jako inspirativní příklad pro výběr počátečních centroidů může sloužit komparativní mezinárodní studie provedená Szokolczaiem a Fustosem [1998], kteří vytvořili klastry v datech pomocí hierarchické metody a následně použili multidimenzionální škálování na objekty v každém shluku k výběru několika skutečně reprezentativních objektů pro

vytvoření centroidů klastrů. Srovnali je navzájem a vybrali počáteční centra ke klastrování celého datového souboru metodou K-průměrů.

Poté se v metodě K-průměrů postupně zkoumají vzdálenosti každého objektu od každého počátečního centroidu tak, že pro každou takovou dvojici spočte algoritmus Euklidovskou vzdálenost a objekt je pak přiřazen k nejbližšímu počátečnímu centroidu. Pro každý vytvořený shluk je spočten nový centroid a opět se postupně zkoumají vzdálenosti každého objektu od každého centroidu. Pokud má objekt blíže k centroidu jiného shluku, je přesunut tam (tzv. relokační mechanismus), toto přearozování objektů nebylo v předešlé hierarchické metodě možné [Garson 2010]. Někdy je objekt přesunut na základě statistické síly struktury sjednocování, spíše než na základě přirozené podobnosti původních objektů, což může následně způsobit spojení výrazně nepodobných prvků [Barker 1976: 44]. Iterace (resp. opakování tohoto procesu) se provádí do doby, dokud se nedojde ke stabilní konfiguraci rozkladu na výsledné shluky. Výzkumník volí maximální počet iterací, ale pokud se proces zastaví až na něm, může se v analýze zvolit větší číslo z důvodu případné nestability řešení.

Je vhodné provést analýzu pro několik dělení (vybrat si několik možných řešení s různým počtem prvotních shluků) a následně určit poměr vnitorskupinové a meziskupinové variability pro všechny analýzy (všechny  $k$ ). Jako nejlepší je určen takový počet shluků  $k$ , při kterém je poměr vnitorskupinové a meziskupinové variability nejmenší. Tento algoritmus minimalizuje vnitorslukový rozptyl a maximalizuje rozptyl mezi klastry jako v procesu ANOVY [Garson 2010]. Hlavním cílem metody K-průměrů je opět nalezení takových skupin objektů, u nichž je podobnost uvnitř shluku co největší.

### 9.2.2 Předpoklady a interpretace metody K-průměrů

Z hlediska interpretace nesou centroidy shluků důležitou informaci o tom, zda průměrná hodnota klastru jasně identifikuje nějakou subskupinu populace. Proto je i pro perspektivu budoucí teoretické klasifikace důležité do interpretace zahrnout demografické charakteristiky jednotlivých klastrů, aby byla nalezená společná skupina jedinců snadněji objasnitelná. V rámci komparativních výzkumů je také nezbytné do této charakteristiky skupin jedinců zahrnout jejich příslušnost ke zkoumaným zemím a kulturám, aby se šance na nalezení systémových proměnných, které je do těchto skupin zařazují, co nejvíce zvýšily.

Finální klasifikační řešení by mělo následně využít diskriminační analýzy ke klasifikaci nezařazených případů do klastrů na základě nejbližších vzdáleností. Tato replikace klasifikační analýzy je nezbytná nejen z důvodu určení vhodného počtu klastrů, ale

také kvůli potvrzení smysluplných výsledků řešení metody K-průměrů na nezávislém vzorku. Diskriminační analýza má v tomto kroku shlukování spíše validizační charakter [DiStefano, Kamphaus 2006: 781].

Tento iterativní přístup neprodukuje hierarchické či nahromaděné shluky, ale oddělené stejně velké klastry, které se nepřekrývají, což jak již bylo řečeno je podstatnou a žádanou vlastností shlukovacích metod pro mezinárodní srovnávání dat. Tyto klastry nejsou vždy kulovitého tvaru, neboť homogenní shluky jsou specialitou hierarchických metod, důvodem může být ale vyžadovaná standardizace proměnných v hierarchickém shlukování [Aldenderfer, Blashfield 1984: 48]. Klastrování pomocí metody K-průměrů umožňuje analýzu velkých datových souborů (klasicky  $> 200$ ), neboť proces nevyžaduje primární výpočet matice vzdáleností každého páru objektů, který shlukovací proceduru značně zpomaluje a činí ji ne tolik vhodnou pro komparativní výzkumy. Ale i když je tato metoda vhodná pro datové soubory s velkým počtem objektů, její řešení je závislé na pořadí objektů v datovém souboru, proto požaduje jejich znáhodnění. S tím se pojí značná časová náročnost.

K-means metoda primárně předpokládá nezávislost proměnných, na což například dvoukroková shlukovací metoda citlivá není. Je také značně senzitivní na odlehlá pozorování. V této souvislosti vzniká další kontroverzní pole pro vědce, neboť se v této oblasti střetávají dva pohledy na ideální řešení. Jeden preferuje odlehlá pozorování z analýzy vyloučit, druhý upřednostňuje zařazení objektu do samostatného klastru. Každé řešení má svá pro a proti a výhoda komplexního navrženého řešení se zařazením všech objektů (i odlehlých) do výsledné struktury shluků vždy stojí proti efektivitě řešení, jejíž primárním cílem je redukovat multidimenzionální prostředí na co nejmenší počet vhodných klastrů a odlehlá pozorování vyloučit [Bailey 1994: 44]. S tím souvisí i pozornost věnovaná chybějícím hodnotám (missing values), které K-means z analýzy automaticky vylučuje, a vzniká teda další otázka jejich případného adekvátního nahrazování (např. průměrem hodnot atp.). Poslední připomínka se týká faktu, že shlukování K-průměrů se používá v případě, kdy datový soubor obsahuje pouze kvantitativní (metrické) spojitě proměnné, navíc tato metoda vykazuje lepší výsledky, pokud jich neanalyzuje příliš velké množství.<sup>25</sup>

Pozitivní vlastnosti pro komparativní výzkumy se skrývají u metody K-průměrů v její schopnosti pracovat již s větším množstvím mezinárodních dat. Umožňuje tím podrobnější analýzu, která se může zaměřit na nalezení důležitých klasifikačních proměnných, které jsou

---

<sup>25</sup> Pro nominální proměnné se používá metoda K-modů, kde je shluk reprezentován vektorem údajů o nejčastěji zastoupené kategorii jednotlivých proměnných. Jako míra nepodobnosti se používá koeficient prosté shody [Řezanková, Húsek, Snášel 2009: 89].



schopny objasnit některé mezinárodní odlišnosti, popřípadě identifikovat, zda jsou tyto klasifikační proměnné spojené se sociálním prostředím zkoumaných zemích, které by vysvětlovalo podstatu podobnosti sdružení stejných zemí. Na druhou stranu má metoda K-průměrů v souvislosti s mezinárodními výzkumy značné nedostatky, neumožňuje kombinovat v analýze vliv spojitých a kategoriálních proměnných a ani není připravena na velké množství analyzovaných proměnných, což komparativní výzkum i pro explorační účely vyžaduje. Je dále citlivá na odlehlá pozorování, která se v mezinárodních datech také často vyskytují a je nutné je správně identifikovat. Největší kritika shlukování K-průměrů se snáší ze strany velkého vlivu subjektivity výzkumníka, který ovlivňuje výběr řešení prvotním zvolením počtu shluků. Dále je kritizován nedostatek statistických indexů při výběru finálního řešení a citlivost shlukovacího řešení vůči klastrovému algoritmu [DiStefano, Kamphaus 2006: 781].

Metoda K-means z těchto důvodů lépe slouží jako vylepšení hierarchických technik. První vlna klastrování může využít hierarchickou metodu a potom centroidy vytvořených klastrů použít v druhé vlně shlukování jako centra shluků pro K-means přístup, který umožňuje přemístění objektů v klastrech a napravuje tak potenciální chyby ve shlukování [Rencher 2002: 482]. Navíc se může metoda K-means projevit i jako kvalitní diagnostický nástroj vzhledem k předešlým výsledkům hierarchické metody, neboť schopností zmapovat průměry proměnných ve finálních klastrech z hierarchické metody ukazuje základní charakter shluků, ve smyslu interpretace jejich klasifikačních proměnných. Přidavkem může potvrdit příslušnost objektů ke zvoleným klastrům z předešlých analýz. Z tohoto úhlu pohledu se jeví metoda K-průměrů jako ideální varianta ve spojení s hierarchickou metodou.

Výstupy metody K-means mají podobu *tabulky* ANOVY, jež testuje signifikantní přínosy klasifikačních proměnných v analýze (viz. Tabulka 9.3). Tabulka *Final Cluster Centers* pak vypovídá o zmíněných průměrech proměnných v jednotlivých klastrech, které mají schopnost vybrané klastry interpretovat (viz. Tabulka 9.4). Posledním numerickým výstupem je tabulka *Cluster membership*, která zařazuje objekty k příslušným shlukům a ukazuje konkrétní vzdálenosti objektů od centra shluku (totožně jako v hierarchické metodě) (viz. Tabulka 9.5).<sup>26</sup>

---

<sup>26</sup> Následující výstupy SPSS pracují se 22 000 validních respondentů ze 22 zemí, spojitě (nezkorelované) proměnné: důvěra v policii a ochota přijmout imigranta jiné rasy do rodiny, data ESS (2002/2003) jsou agregována podle země. Otázka pro shlukovou analýzu je položena tímto způsobem – jak se respondenti, všichni občané zařazených zemí ve výzkumu ESS (2002), liší svou ochotou přijmout přistěhovalce jiné rasy do své rodiny a vyjadřovanou důvěrou v policii dané země v jednotlivých zemích.

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
imdetmr_mean imigranti jiné rasy_sňatek_škála blízkosti	4,501	3	,287	18	15,681	,000
trstplc_mean důvěra v policii	4,193	3	,225	18	18,628	,000

Tabulka 9.3. Tabulka ANOVY (test signifikance proměnných)

Tabulka ANOVY testuje signifikantní přínosy klasifikačních proměnných ve shlukovací analýze metody K-průměrů. Obě dvě vybrané nezkorelované proměnné z mezinárodních dat ESS prokázaly signifikantní přínos pro tvorbu výsledných shluků. Pokud by tyto proměnné testem ANOVY neprošly (neprokázaly by přínos do shlukování), neměly by se brát v dalších procesech analýzy v potaz a vyřadit se, nijak by totiž od sebe shluky jedinců dobře neodlišovaly. Test signifikance je velkým přínosem metody K-průměrů do analýzy mezinárodních dat, neboť pokud výzkumníci mají zájem interpretovat nalezené shluky a nalézat proměnné na systémové úrovni, je tento údaj nezbytný.

Final Cluster Centers				
	Cluster			
	1	2	3	4
imdetmr_mean imigranti jiné rasy_sňatek_škála blízkosti	2,57	3,78	5,61	3,61
trstplc_mean důvěra v policii	6,60	7,20	6,31	5,37

Tabulka 9.4. Final Cluster Centers (průměry proměnných v klastrech)

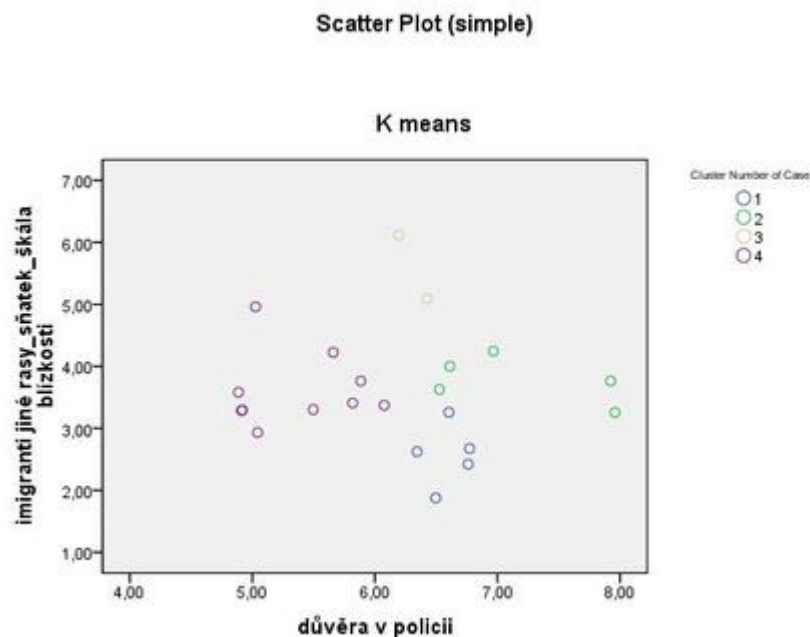
Tabulka *Final Cluster Centers* vypovídá o průměrech proměnných v jednotlivých klastrech, které mají schopnost vybrané shluky interpretovat. Na počátku analýzy byly zvoleny počáteční 4 klastry, které od sebe vybrané spojitě proměnné v mezinárodních datech ESS odlišily následujícím způsobem. Například čtvrtý shluk zemí se vyznačuje nejnižší důvěrou v policii, ale relativně vyšší ochotou přijmout do rodiny přistěhovalce jiné rasy. Třetí shluk je naproti tomu charakteristický nejnižší ochotou přijmout přistěhovalce jiné rasy do rodiny, ale důvěra v policii má v tomto shluku průměrnou až vyšší hodnotu. Vidíme, že ani metoda K-průměrů nemá tak rozsáhlé interpretační možnosti ohledně vysvětlení vlivu proměnných na shluky různých zemí. Narozdíl od hierarchických metod se však tyto možnosti výrazně zlepšily, tato metoda dokáže shluky charakterizovat a má informace o signifikanci proměnných. Pokud bychom navíc do analýzy přidali více spojitých proměnných, interpretační schopnost metody K-průměrů by se ještě více zvýšila společně s pravděpodobností nalezení důležitých systémových proměnných nebo skrytých proměnných, ale i tyto možnosti jsou poněkud omezené, neboť metoda K-průměrů lépe funguje při menším počtu proměnných. Vůči požadavkům komparativních výzkumů tato technika postrádá schopnost pracovat s kategoriálními proměnnými a nedokáže identifikovat skutečný přínos spojitých proměnných do vytvoření klastřů. Stále je patrný spíše explorační potenciál, kterým mohou shlukovací metody vytvořit přehled ve velkém množství dat a najít potenciální strukturu, která se v nich může skrývat.

Ca...	cntry Country	Cluster	Distance
1	Austria	1	,701
2	Belgium	4	,677
3	Czech Republic	4	1,392
4	Germany	1	,686
5	Denmark	2	,727
6	Spain	4	,334
7	Finland	2	,924
8	Francia	4	,535
9	Unired Kingdom	4	,744
10	Greece	3	,528
11	Hungary	4	,564
12	Switzerland	1	,205
13	Ireland	2	,687
14	Israel	3	,528
15	Italy	2	,627
16	Luxembourg	1	,257
17	Neitherland	4	,491
18	Norway	2	,520
19	Poland	4	,555
20	Portugal	4	,756
21	Sweden	1	,223
22	Slovenia	4	,487

Tabulka 9.5. Cluster membership (vzdálenosti objektů od shluku)

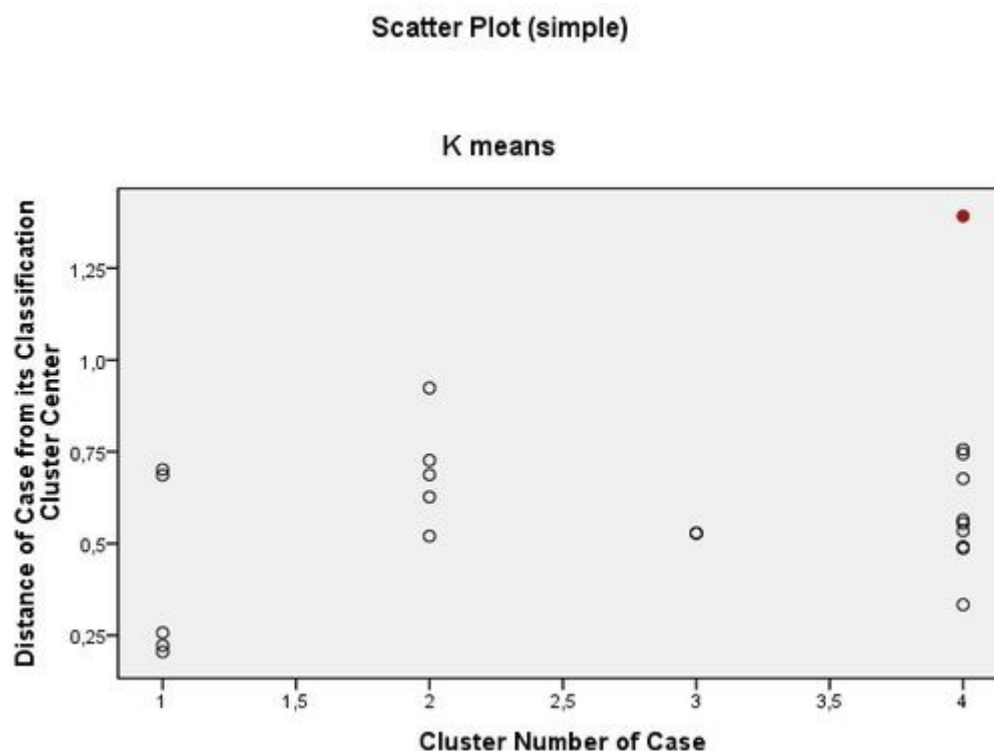
Tabulka *Cluster membership* zařazuje objekty zemí z mezinárodního výzkumu k příslušným shlukům, ke kterým patří. V tuto chvíli lze přesně zjistit, jaké země náleží do zmíněného čtvrtého klastu, který se vyznačoval nejnižší důvěrou v policii a relativně vyšší ochotou přijmout do rodiny přistěhovalce jiné rasy, jsou to země převážně střední, východní a jižní Evropy. Na rozdíl od výsledků Wardovy hierarchické metody jsme do analýzy K-průměrů zvolili 4 klastry, proto je rozdělení zemí poněkud jiné než v předešlé metodě, i když metody pracují se stejným vzorkem respondentů a stejnými proměnnými. Graficky jsou tyto rozdíly lépe patrné ze Scatter Plotu (Graf 9.15). Dále tato tabulka ukazuje konkrétní vzdálenosti objektů zemí od centra shluku a identifikuje tak odlehlá pozorování, nejodlehlejším pozorováním ve čtvrtém klastu je Česká republika (tyto vzdálenosti jsou vizuálně lépe znázorněny ve Scatter plotu vzdáleností – viz. Graf 9.16).

Grafické možnosti výstupů metody K-means jsou značně omezené, lze vytvořit již jednou zmíněný *Scatter Plot* objektů pro dvě proměnné (viz. Graf 9.15) a *Scatter Plot pro ilustraci vzdáleností* objektů od centra daného shluku (viz. Graf 9.16).



Graf 9.15. Scatter Plot K-průměrů (dvě proměnné)

*Scatter Plot* je grafickým znázorněním výsledků vypsanych v tabulce Final Cluster Centers (viz. Tabulka 9.4). Lze z něj vyčíst zcela stejné výsledky o podobě klastru vzhledem k průměrným hodnotám klasifikačních proměnných, jen je vizuálně přehlednější a podává lepší shrnutí vzájemného postavení výsledných shluků mezi sebou. Také ukazuje, jak jsou země rozptýleny v rámci jednoho shluku a zda jsou klastry jasně oddělené a nepřekrývají se. Pokud bychom do analýzy K-průměrů zařadili více proměnných, vznikla by nám série dvourozměrných Scatter plotů, ze kterých lze vyčíst charakter shluků v různých dimenzích jednotlivých klasifikačních proměnných. Jinou možností by bylo vytvořit pro tři proměnné Scatter Plot ve 3D verzi, ale v tomto případě se názornost oddělených shluků ztrácí a interpretace je složitější. Rozptyl zemí (objektů) uvnitř jednoho shluku lépe zachycují příklady *Scatter Plotů vzdáleností objektů od středu shluku*, umístěné níže (viz. Graf 9.16). Tyto Scatter Ploty rozptyl shluků ukazují klasickým způsobem, na ose X jsou umístěny shluky, vertikálně jsou viditelné rozptýlené objekty. Čtvrtý shluk je typický nejvíce rozptýlenými zeměmi, lze o něm tedy říci, že je nejméně homogenní a nejméně soudržný. Země, které do něj spadají, se ve svých názorech na důvěru v polici a ochotou přijmout přistěhovalce jiné rasy do rodiny (v rámci jednoho shluku) nejvíce odlišují. Lze z tohoto výstupu identifikovat přítomnost odlehlých pozorování, ale nelze je blíže specifikovat, pouze víme, že by mohly potenciálně ohrozit celkové výsledky metody K-průměrů. Nalézt konkrétní země s největší vzdáleností od středu shluku je možné jen podle numerických výstupů Cluster membership (viz. Tabulka 9.5), tak je možné například zjistit, že nejodlehlejší zemí ve čtvrtém klastru je Česká Republika.



Graf 9.16. Scatter Plot K-průměrů (vzdálenosti objektů od centra shluku), odlehlé pozorování

### 9.3 Dvoukroková shluková analýza

Procedura dvoukrokové shlukové analýzy je statistický explorační nástroj k odhalení přirozeného shlukování uvnitř datového souboru, které není na první pohled patrné. Používaný algoritmus dvoukrokové shlukové analýzy má několik žádoucích vlastností, které jej odlišují od tradičních klastrovacích metod [Garson 2010]:

- ◆ schopnost nalézt klastry na základě *kombinace kategoriálních a spojitých proměnných*, již při nejméně jedné kategoriální proměnné se doporučuje provést dvoukrokovou klastrovou analýzu (ostatní se používají na spojitě a skutečně binární proměnné). Klastrovací mechanismus dvoukrokové shlukové analýzy pracuje se standardizovanými spojitými proměnnými, pokud jakákoli spojitá proměnná není ještě standardizovaná, musí být umístěna do tzv. standardizačního listu (Standardized list).
- ◆ *automatická selekce počtu klastrů*
- ◆ schopnost efektivně analyzovat *velký datový soubor*

Všechny vyjmenované vlastnosti činí z dvoukrokové shlukové analýzy neocenitelný přínos pro cíle komparativních výzkumů. Dokáže se výpočetně vypořádat s velkým množstvím mezinárodních dat, které je žádoucí v komparativních studiích analyzovat a nacházet pro jejich vysvětlení systémové proměnné, které dokáží za název země dosadit relevantní strukturní proměnnou. Navíc má schopnost v analýzách kombinovat velké množství jak spojitých, tak i kategoriálních proměnných (resp. dichotomických proměnných), a tím pádem se výrazně zvyšuje šance jednak na kvalitní interpretaci nalezených shluků, včetně jejich základních demografických znaků, a také na nalezení žádoucích systémových proměnných, které mohou být vkladem do nově vznikající teorie. Ve dvoukrokové shlukové analýze odpadají i starosti se subjektivním ovlivněním výsledného řešení výzkumníkem, neboť v procesu této metody se již nemusí předem volit optimální počty klastrů, mechanismus dvoukrokové analýzy je nalezne a vyhodnotí sám (jde o tzv. objektivní přístup ke shlukování). I když je i dvoukroková shluková analýza stále spíše exploračním nástrojem v analýze mezinárodních dat, její výsledky již lze rozumně zvážit a připravit na další hlubší analýzu jinými multidimenzionálními technikami.

Představuje pomyslný předěl mezi nejmodernějšími shlukovacími technikami a tradičními shlukovacími metodami (viz. hierarchická metoda a metoda K-průměrů), je totiž variantou moderní modifikace hierarchických metod, tudíž dokáže svými pokročilejšími výstupy demonstrovat, jaké vylepšení lze v analýze mezinárodních dat od moderních shlukovacích metod očekávat. Navíc je daleko lépe dostupná v uživatelských statistických

programech (např. v SPSS) než zmíněné verze moderních vícerozměrných shlukovacích technik.

### 9.3.1 Postup shlukování ve dvoukrokové shlukovací metodě

Dvoukroková klastrovací analýza je rozšířením práce Banfielda a Raftery (1993), kteří rozvíjeli klastrovací metodu pro spojitě proměnné, založenou na redukci věrohodnostní míry při spojování klastrů. A také vychází z práce Melia a Heckermana (1998), používající podobný pravděpodobnostní přístup pro kategoriální proměnné [podrobněji Garson 2010]. Metoda dvoukrokové shlukové analýzy byla softwarově implementována do systému SPSS v podobě procedury *TwoStep Cluster Analysis*, rozvíjenou na klastrovacím algoritmu BIRCH. Datovým souborem se prochází pouze jednou (one-pass-through-the-data approach). Pokud se v určitém řádku vstupní matice vyskytuje chybějící údaj, je celý řádek (objekt-respondent) vynechán [Řezanková, Húsek, Snášel 2009: 109].

Algoritmus je realizován ve dvou fázích. V první fázi se objekty shlukují do malých shluků (podshluků), jejichž počet je podstatně menší než počet objektů původního souboru. Objekty se v pořadí daném datovým souborem postupně hodnotí, zda mohou být přímo zařazeny do již vytvořeného shluku, nebo zda bude vytvořen nový shluk, je proto vhodné, aby byly objekty náhodně uspořádány.

Vytváří se modifikovaný tzv. CF-strom. Jedná se o klasické zařazování objektů do vytvořeného shlukovacího stromu (příklad dendrogramu viz. Graf 9.11), který lze graficky zobrazit jako schéma, jehož prvky jsou uzly a větve na různých úrovních. Nejnížší úrovně těchto uzlů zahrnují v tomto případě podshluky objektů, vytvořené v první fázi. Každý vstup do uzlu je popsán *charakteristikou CF* (Cluster Features), která se skládá z počtu objektů vstupu, střední hodnoty a rozptylu každé spojitě proměnné a četností každé kategorie každé kategoriální proměnné. Pokud se objekt při shlukování nachází v rámci prahové vzdálenosti od nejbližšího podshluku, vstupuje do tohoto podshluku. V opačném případě daný objekt vytváří vlastní podshluk. Není-li již v příslušném uzlu prostor pro zařazení dalšího vstupu, je tento podshluk rozdělen na dva, a to podle nejvzdálenějších vstupů. Ostatní vstupy se slučují na základě kritéria blízkosti. Pokud CF-strom roste za povolenou maximální velikost, je strom přestavěn zvětšením prahové vzdálenosti objektů [Řezanková, Húsek, Snášel 2009: 110].

Všechny objekty určitého vstupu jsou reprezentovány souhrnně pomocí CF charakteristiky. Po zařazení nového objektu do vstupu se tato charakteristika přepočítává (aktualizuje), a to na základě nového objektu a původní CF charakteristiky. Součástí

algoritmu může být zjišťování *odlehlych objektů*, tj. těch, které se nehodí do žádného shluku. Je-li počet objektů ve vstupu menší než stanovená velikost toho největšího uzlu v CF-stromu, jsou tyto objekty považovány za odlehlé. CF–strom je přestavěn bez těchto objektů, a poté se zkoumá, zda by mohly být znovu odlehlé objekty zařazeny. Pokud nejsou odlehlá pozorování umístěna do nových klastrů, jsou vyřazena. Pokud se výzkumník nezaměří na zjišťování odlehlých pozorování, po finálním klastrování jsou nezařazená pozorování označena jako odlehlá (outliers). Odlehlému klastru je dáno číslo -1 pro identifikaci a není zařazen do výsledného počtu klastrů [Garson 2010] (viz. Tabulka 9.6). Výsledkem první fáze algoritmu je nová datová matice, ve které jsou v řádcích charakteristiky jednotlivých podshluků.<sup>27</sup>

	N	% of Combined	% of Total
Cluster 1	13789	41,2%	33,9%
2	4217	12,6%	10,4%
3	12265	36,6%	30,2%
4	3083	9,2%	7,6%
Outlier (-1)	155	,5%	,4%
Combined	33509	100,0%	82,4%
Excluded Cases	7137		17,6%
Total	40646		100,0%

Tabulka 9.6. Cluster Distribution (příklad odlehlého klastru)

Pozornost, která je ve dvoukrokové shlukové analýze věnována odlehlým pozorováním, je vítanou vlastností pro komparativní výzkumy, protože odlehlá pozorování jsou častým problémem, se kterým se musí v analýze mezinárodních dat výzkumníci potýkat. Cení se jak schopnost dvoukrokové analýzy opětovně odlehlé objekty zařazovat do existujících shluků, tak i možnost vydělit tato pozorování do speciálně označeného klastru, u něhož lze také zjišťovat nejružnější charakteristiky, velikost a interpretovat ho na základě zahrnutých klasifikačních proměnných. Tím, že se dvoukroková analýza soustředí na identifikaci odlehlých objektů, mají tato pozorování menší možnost ovlivnit výsledky shlukování. Toto numerické shrnutí relativního a absolutního množství objektů v každém shluku v tabulce Cluster Distribution umožňuje dvoukrokovou shlukovou analýzu i v grafické podobě koláčového grafu, kde je velikost shluků názornější, největší je zde první klast jedinců.

Ve druhé fázi algoritmu je každý vytvořený podshluk přiřazen do některého

<sup>27</sup> Následující výstupy SPSS pracují se 33 509 validními respondenty ze 22 zemí, data ESS (2002/2003) už nejsou agregována podle zemí. KATEGORIÁLNÍ PROMĚNNÉ: země, pohlaví, přátelé z jiné země, pracovní kolegové z jiné země, povolit vstup do země lidem stejné rasy/jiné rasy, lidem z bohatších zemí a chudších zemí; SPOJITÉ PROMĚNNÉ: ochota přijmout do rodiny přistěhovalce jiné rasy, názory na ovlivnění ekonomiky země, zločinnosti v zemi a obohacení kultury imigrantstvím. Otázka pro shlukovou analýzu - jak se respondenti, všichni občané zařazených zemí ve výzkumu ESS (2002), liší mezi sebou ve všech zahrnutých proměnných.

z konečných shluků, jejichž počet je předem stanoven. Protože počet podshluků je podstatně menší než počet objektů původního souboru, mohou být již využity tradiční metody shlukování. Tato fáze se realizuje pomocí hierarchické shlukové analýzy.

Pro stanovení konečného optimálního počtu shluků lze využít tzv. informační kritéria [Řezanková, Húsek, Snášel 2009: 132]. Tyto indexy jsou kritéria používaná ke srovnávání různých modelových řešení a k determinaci nejlepšího modelu, který vyhovuje datovému souboru; říkají nakolik přidaný shluk vylepšuje konkrétní model [DiStefano, Kamphaus 2006]. Rozlišují se dvě tato kritéria, přičemž jedno je založeno na bayesovském přístupu, to bývá označováno zkratkou *BIC* (Bayesian Information Criterion) a nazývá se Schwarzovo bayesovské kritérium. Druhé kritérium je Akaikeho, označuje se *AIC* (Akaike Information Criterion). Obě dvě uvedená kritéria mohou být použita v proceduře dvoukrokové shlukové analýzy.

Nejdříve jsou spočteny hodnoty BIC, resp. AIC, pro každý potenciální počet shluků v rámci specifikovaného intervalu. Na jejich základě se stanoví počáteční odhad pro počet shluků, který odpovídá minimu ze zjištěných hodnot. Jinými slovy čím menší hodnota BIC, tím lepší klastrový model se výzkumníkovi nabízí. Poté je počáteční odhad počtu shluků upraven zjišťováním největšího vzrůstu ve vzdálenosti mezi dvěma nejbližšími shluky v každé etapě hierarchického shlukování. V této situaci je změna BIC a změna ve vzdálenostech shluků ohodnocena jako pravý determinant nejlepšího klastrového řešení (největší *Ratio of Distance Measures* – viz. Tabulka 9.7). Počet výsledných klastrů lze opět specifikovat ručně, například na základě předchozího výsledku v hierarchických metodách [Garson 2010].

Num ber of ...	Schwarz's Bayesian Criterion (BIC)	BIC Change <sup>a</sup>	Ratio of BIC Changes <sup>b</sup>	Ratio of Distance Measures <sup>c</sup>
1	424372,757			
2	362409,658	-61963,099	1,000	2,248
3	335103,482	-27306,176	,441	1,214
4	312689,649	-22 413,83	0,36	2,73
5	304778,121	-7911,528	,128	1,154
6	297982,263	-6795,858	,110	1,237
7	292577,511	-5404,752	,087	1,028
8	287332,267	-5245,244	,085	1,203
9	283048,948	-4283,319	,069	1,420
10	280167,976	-2880,972	,046	1,057
11	277467,689	-2700,287	,044	1,120
12	275105,411	-2362,277	,038	1,038
13	272845,923	-2259,489	,036	<sup>d</sup>

Tabulka 9.7. Tabulka Auto-Clustering: informační kritérium BIC

Dvoukroková shluková analýza mezinárodních dat ESS ohodnotila jako nejlepší řešení vytvoření 4 klastrů mezi všemi zahrnutými jedinci do analýzy, na základě největší hodnoty *Ratio of Distance Measures* v tabulce Auto-clustering, která znázorňuje změny hodnot informačního kritéria BIC (odlehlá pozorování se do analýzy nezahrnují).



V obou fázích shlukování se používá stejná míra nepodobnosti pro měření vzdálenosti mezi shluky, podshluky a objekty. Na výběr je standardně ze dvou typů měř (např. v SPSS), a to *Euklidovské vzdálenosti* a *Log-likelihood*. Euklidovská vzdálenost může být použita pouze v případě, že všechny proměnné v analýze jsou kvantitativní spojité. Log-likelihood je vhodná pro kategoriální proměnné, může být ale použita též pro proměnné spojité, případně pro datový soubor obsahující proměnné obou typů. Z tohoto důvodu se většinou v analýzách upřednostňuje, protože vytváření kombinací různých druhů proměnných je pro komparativní výzkumy žádoucí. Vzdálenost mezi dvěma shluky zohledňuje pokles Log-likelihood, jenž nastává při spojení dvou shluků do jednoho.

Předpokladem pro použití této míry je normální rozdělení (Gaussian distribution) pro spojité proměnné a multinomické rozdělení pro kategoriální proměnné. Dále se předpokládá nezávislost pro všechny dvojice proměnných [Řezanková, Húsek, Snášel 2009: 111]. Empirické testování ale ukazuje, že procedura dvoukrokové analýzy (i Log-likelihood) jsou značně robustní vůči porušení všech předpokladů, ale výzkumník si musí být vědom, nakolik jsou tyto předpoklady naplněny [Garson 2010].<sup>28</sup>

### 9.3.2 Interpretace výstupů dvoukrokové shlukové analýzy

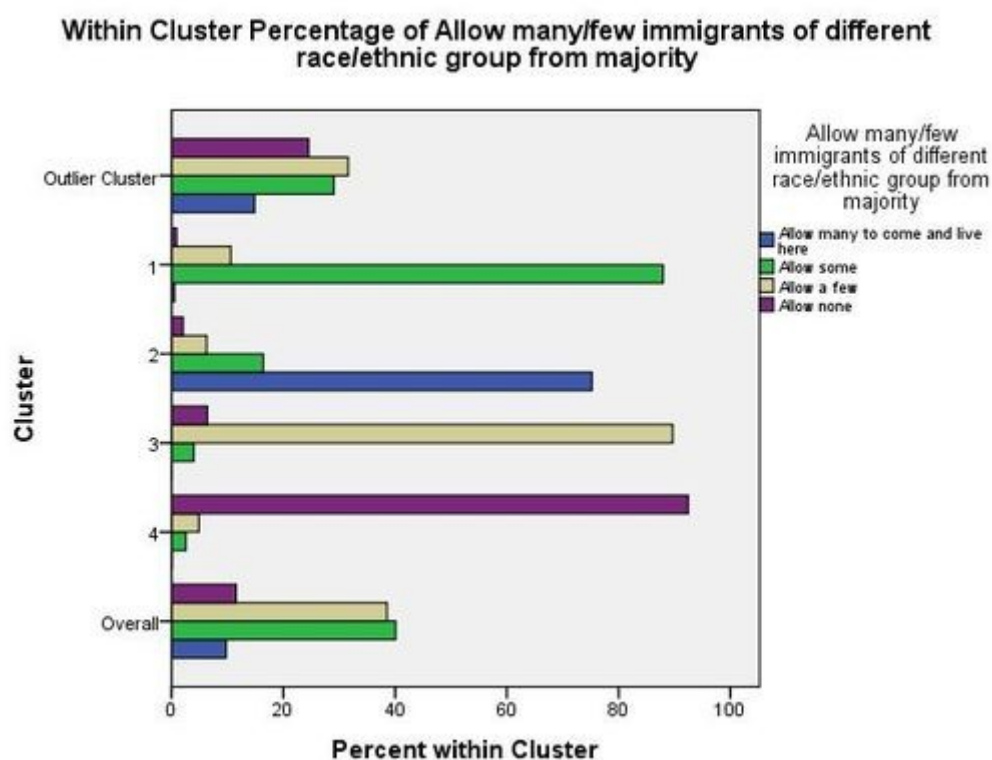
Vedle klasických numerických výstupů v podobě profilů klastrů z průměrných hodnot spojitých proměnných a frekvenčních tabulek pro kategoriální proměnné disponuje dvoukroková shluková analýza širokými možnostmi v tvorbě grafů, na rozdíl od předešlé hierarchické metody a metody K-průměrů. Pro komparativní studie to znamená, že se již konečně otevírá pole možností pro názornou a kvalitní interpretaci nalezených shluků, včetně mapování demografických znaků (pokud budou do analýzy zahrnuty). Také se zvyšuje jejich šance nalézt hledané vysvětlující proměnné propojené se sociálním prostředím, které by objasnilo vznik shluků stejných zemí nebo skupin stejných jedinců z různých zemí. Všechny tyto nalezené položky pak mohou být užitečným vkladem do nové teorie či vznikající klasifikace a nebo podkladem pro podrobnější analýzu jiných multidimenzionálních technik. Následující výstupy nesuplují interpretační hledání těchto strukturních proměnných, jen ukazují explorační sílu pokrokovější shlukové techniky na mezinárodních datech, která již může být efektivním startem pro jiné druhy analýz. Dvoukroková shluková analýza využívá

---

<sup>28</sup> Např. ve výstupech mezinárodních dat ESS (2002) pro dvoukrokovou shlukovou analýzu je vhodné zmínit, že všechny spojité proměnné jsou více méně zkorelované (nejvíce proměnné přínosy imigrantství do ekonomiky země a obohacení kulturního života,  $r = 0,56$ ), N rozdělení mají pouze proměnné přínosy imigrantství do ekonomiky země a obohacení kulturního života, ostatní spojité proměnné mají sešikmené rozdělení.

grafické možnosti ve vytváření *sloupcových diagramů klastrových frekvencí* (bar charts of cluster frequencies), *koláčových grafů klastrových frekvencí* (pie charts of cluster frequencies) a *grafů důležitosti proměnných* (variable importance charts).

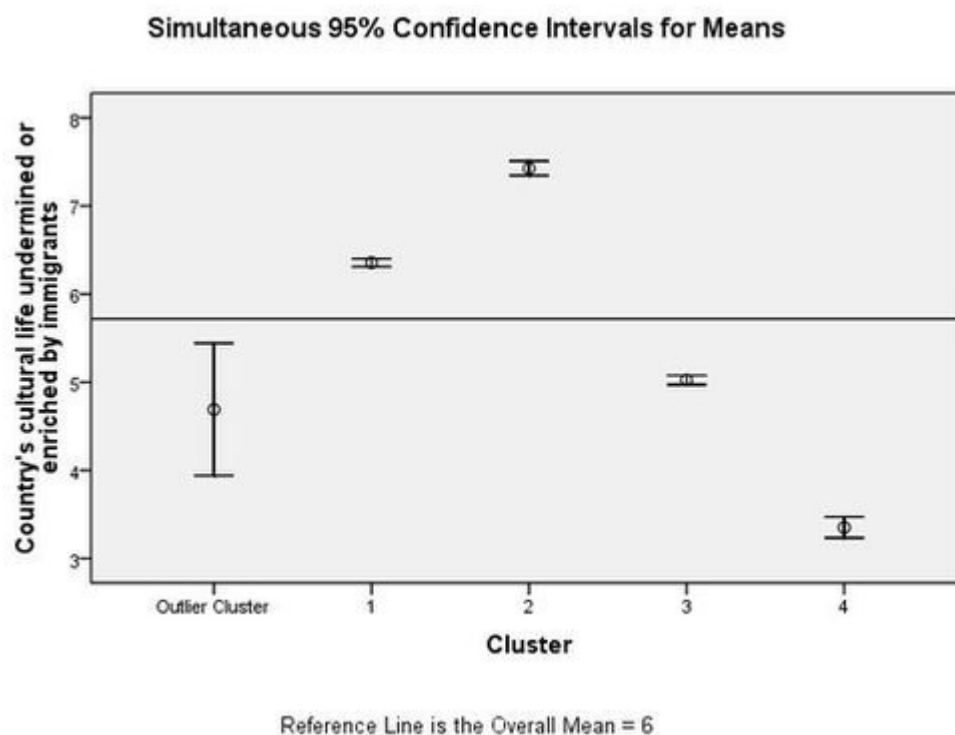
Pro kategoriální proměnné se standardně vytvářejí sloupcové diagramy znázorňující frekvenci všech kategorií proměnných v jednotlivých klastrech (viz. Graf 9.17). Pro každou spojitou proměnnou je k dispozici také sloupcový diagram ukazující rozptyl spojité proměnné uvnitř klastru, tzv. chybové složky klastrů (viz. Graf 9.18). Koláčový graf jednoduše znázorňuje relativní množství respondentů uvnitř každého klastru.



Graf 9.17. Sloupcový diagram relativních četností kategorií pro kategoriální proměnnou

Tento *sloupcový diagram* je vytvořený pro kategoriální proměnnou, která říká, do jaké míry by podle respondentů z výzkumu ESS měla země, již jsou občany, povolit vstup do země přistěhovalcům jiné rasy či etnické skupiny než je většina žijících lidí v dané zemi. Kategorie, které jsou vyhodnoceny, jsou: dovolit vstup hodně přistěhovalcům, jen některým, malému počtu a nikomu. Rozložení četností jednotlivých kategorií proměnné vidíme v jednotlivých klastrech, v klastru odlehlých pozorování i v celkovém součtu. Například první klastr se vyznačuje tím, že jedinci v něm zařazení si myslí, že by některým přistěhovalcům jiné rasy či etnické skupiny vstup do země měl být povolen. Právě toto přehledné

znázornění umožňuje výzkumníkovi relativně přesně a rychle interpretovat povahu jednotlivých klastrů a smysluplné rozdělení jednotlivců do různých skupin podle vybraných proměnných. Vidíme, že interpretační potenciál shlukové analýzy významně vzrůstá na rozdíl od možností předchozích tradičních shlukovacích metod. V tomto typu analýzy si už výzkumníci skutečně mohou udělat představu o charakteru struktury velkého množství mezinárodních dat.



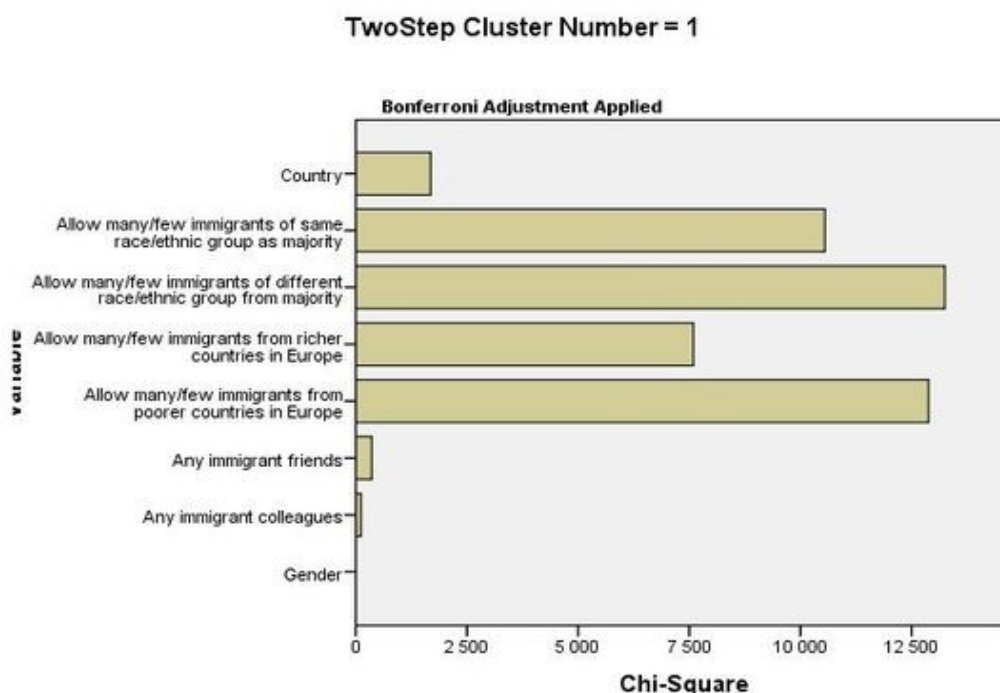
Graf 9.18. Sloupcový diagram vnitroklastrového rozptylu pro spojitou proměnnou

Tento sloupcový diagram je vytvořený pro spojitou proměnnou, která říká, zda si respondenti (občané země) z mezinárodního výzkumu ESS myslí, že lidé z jiných zemí, kteří přicházejí do země žít, většinou narušují kulturu země nebo ji obohacují. U spojitou proměnnou nemohou být vyhodnocovány jejich kategorie, proto graf ukazuje rozptyl spojitou proměnnou uvnitř klastru, tzv. chybové složky klastrů v podobě box plotů, a umístění klastru vzhledem k průměrné hodnotě, kterou zvolená proměnná ve výzkumů celkově nabývá. Z výstupu se zdá, že největší rozptyl má čtvrtý klaster (pokud nepočítáme klaster odlehlých pozorování) a například první klaster zaujímá druhé nejkladnější stanovisko vůči přínosu imigrantství do místní kultury. Lze takto porovnávat výsledné shluky mezi sebou, v rámci všech klasifikačních proměnných.

Grafy důležitosti proměnných (variable importance charts) jsou pro shlukovou dvoukrokovou analýzu v podstatě nejdůležitější, protože dokáží seřadit všechny klasifikované proměnné (zvláště spojité a zvláště kategoriální) v každém klastru postupně podle jejich důležitosti, resp. podle míry přispění k rozdělení objektů do výsledného shluku. Pro určení důležitosti příspěvku proměnných k vytvoření klastru se pro *kategoriální proměnné* používá Chi-square nebo t-test signifikance (resp. statistika Pearson chi-square) (viz. Graf 9.19) a jako ukazatel příspěvku *spojitých proměnných* se používá T-statistika (shoda průměrů), negativní hodnota T-statistiky znamená, že proměnná má menší hodnotu než je její průměrná hodnota ve výzkumu (viz. Graf 9.20).

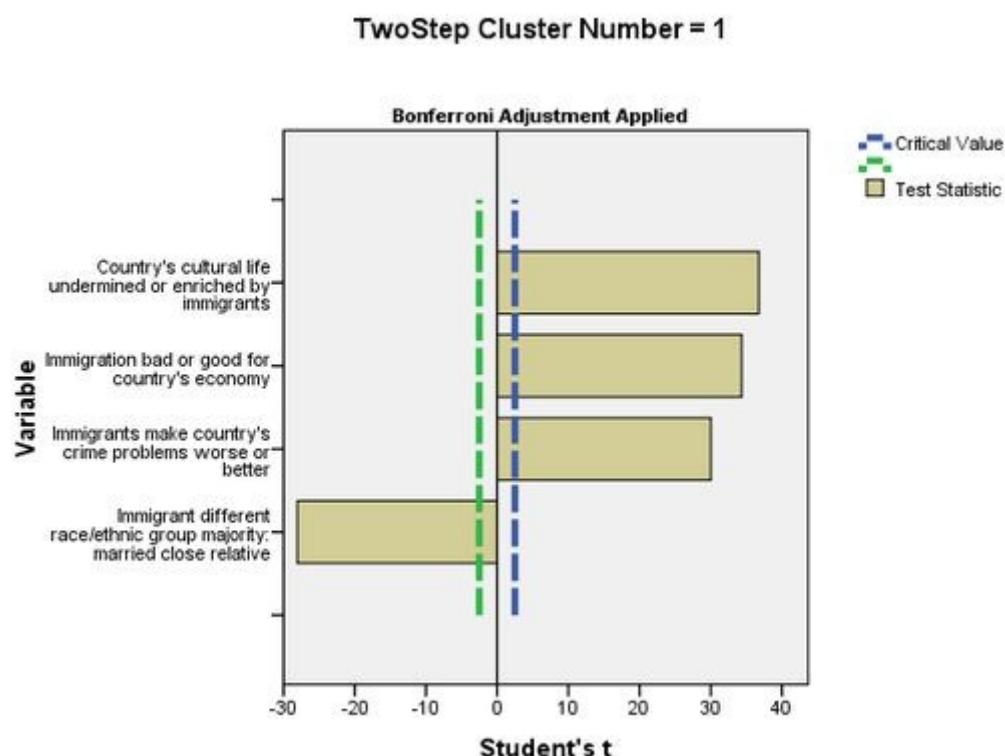
Grafy důležitosti jsou obzvláště podnětné pro komparativní analýzy, neboť na základě nich lze rozhodnout o tom, zda se bude klasifikační proměnná zvažovat do dalších typů multidimenzionálních analýz. Mapovat lze důležitost všech proměnných pro každý jednotlivý klastr (viz. Graf 9.20) nebo je možné graficky znázornit důležitost jedné proměnné pro tvorbu všech klastrů (viz. Graf 9.21). Ty proměnné, u nichž se identifikuje malý podíl na tvorbě výsledných klastrů, se po řádném zvážení mohou z analýzy vyloučit jako nesignifikantní a ani nezvažovat v analýze mezinárodních dat dál. Na základě těchto údajů lze také modifikovat interpretaci shluků, neboť výsledky sloupcových grafů o významu jednotlivých proměnných pro výsledné shluky by měly být konfrontovány právě se zjištěnou důležitostí dané proměnné. Pokud proměnná není příliš důležitá pro vytvořený shluk, ani její interpretace není dále přínosná.

V informaci, kterou přinášejí grafy důležitosti proměnných, spočívá také největší rozdíl mezi předešlými tradičními shlukovacími metodami a dvoukrokovou shlukovací metodou. V hierarchické aglomerativní metodě jsme byli schopni vypátrat pouze hierarchickou strukturu shlukování a došli jsme k omezené interpretaci zvolených výsledných shluků z analýzy. Metoda K-průměrů již byla schopna tuto interpretaci prohloubit a dokázala určit, zda jsou vybrané proměnné pro shlukování signifikantní či nikoli. Dvoukroková shluková analýza přispěla k interpretaci shluků nejvíce, zvláště možnostmi zachycení kombinace vlivu kategoriálních a spojitých proměnných, a navíc kromě potvrzení signifikance jednotlivých proměnných umožnila zjistit, jaký reálný přínos jednotlivé klasifikační proměnné do procesu shlukování přinášejí.



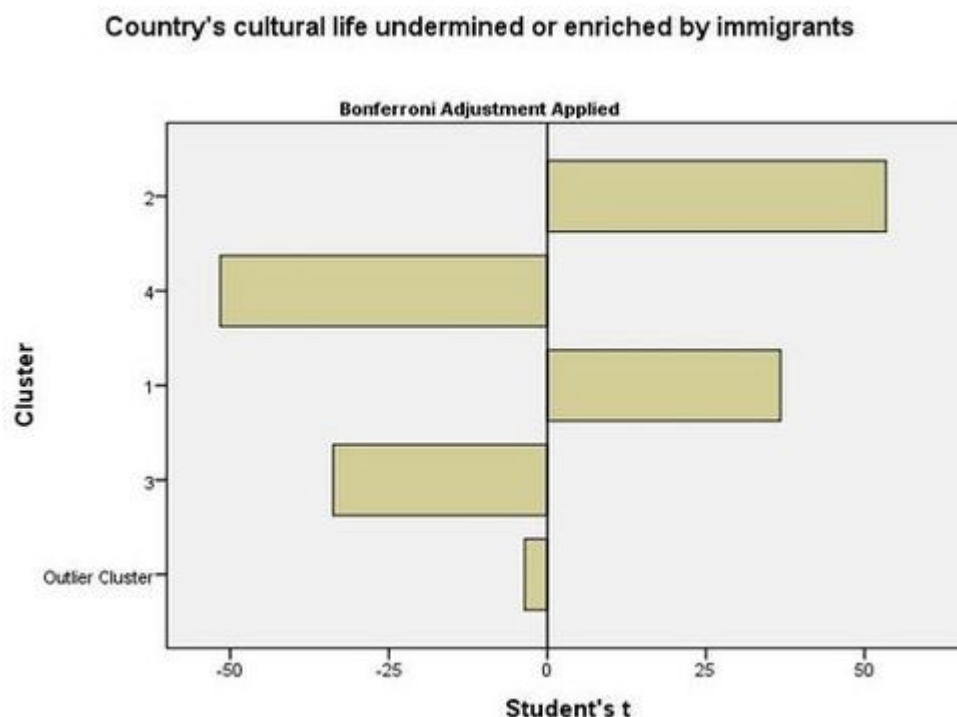
Graf 9.19. Důležitost přínosu kategoriálních proměnných do klastru 1

Graf důležitosti všech kategoriálních proměnných je vytvořen pro první klastr, u kterého jsme již zjistili, že je největší, zaujímá druhé nejkladnější stanovisko vůči přínosu imigrantství do místní kultury a vyznačuje se tím, že jedinci v něm zařazení si myslí, že by některým přistěhovalcům jiné rasy či etnické skupiny vstup do země měl být povolen. V této analýze důležitosti jednotlivých kategoriálních proměnných vidíme, že proměnná o povolení vstupu přistěhovalcům jiné rasy na území státu přináší největší přínos ze všech kategoriálních proměnných k rozdělení tohoto klastru od všech ostatních shluků a v analýze mezinárodních dat bychom jí měli i nadále věnovat pozornost. Z grafu je patrné, že i příslušnost jedinců ke konkrétní zemi je z hlediska tvorby shluků důležitá. Úkolem další analýzy by měla být identifikace rozdělení konkrétních zemí do shluků na základě nejdůležitějších proměnných (jejichž přínos ve dvoukrokové metodě identifikujeme), když jsme již zjistili, že proměnná země má signifikantní přínos. Ty země, které si budou v klasifikačních proměnných blízké, lze pak na Scatter Plotu nalézt v jednom shluku. Naproti tomu například proměnná pohlaví významný přínos do tvorby prvního shluku nemá, na tomto místě by bylo vhodné zjitit její přínos i pro ostatní shluky a v případě, že se bude jevit jako nesignifikantní v celém shlukování, zvážit její přínos pro další analýzu mezinárodních dat.



Graf 9.20. Důležitost přínosu spojených proměnných do klastru 1; interval spolehlivosti

Graf důležitosti všech spojených proměnných je také vytvořen pro první shluk, o němž víme, že je největší, zaujímá druhé nejkladnější stanovisko vůči přínosu imigrantství do místní kultury a vyznačuje se tím, že jedinci v něm zařazení si myslí, že by některým přistěhovalcům jiné rasy či etnické skupiny vstup do země měl být povolen. Opět jsme z analýzy zjistili, že spojitá proměnná přínosu imigrantství do kultury země nese ze všech spojených proměnných v analýze největší přínos pro vytvoření tohoto shluku. Spojitá proměnná, která vypovídá o ochotě přijmout imigranta jiné rasy do rodiny, má negativní hodnotu T statistiky, protože její průměrná hodnota v prvním klastru je menší než její průměrná hodnota v celém výzkumu (v Grafu 9.18 by byla tato proměnná pod čarou, která značí její průměr v celém výzkumu). Interpretačně to znamená, že jedinci v prvním klastru jsou ochotnější přijmout do rodiny přistěhovalce jiné rasy, než je průměrná hodnota ochoty v celém výzkumu u všech jedinců zahrnutých do mezinárodní analýzy. Nízké hodnoty na bodové škále této proměnné totiž znamenají zvyšující se ochotu přijmout imigranty jiné rasy do rodiny. Na základě všech interpretací proměnných v prvním shluku lze jasně říci, že jedinci v něm zahrnutí mají obecně kladný postoj k přistěhovalcům a jejich vlivu na různé aspekty domovské země. V grafu lze také na vertikální linii identifikovat interval spolehlivosti, jež určuje, které proměnné by jako nesignifikantní již dále analýzou neprošly (hraje roli kritické hodnoty testu shody distribucí proměnných uvnitř klastru). Vidíme, že v prvním klastru jsou všechny spojitě proměnné signifikantní.



Graf 9.21. Důležitost přínosu jedné spojité proměnné do všech klastrů

Toto je *graf důležitosti* vytvořený pro jednu spojitou proměnnou, která říká, zda si respondenti (občané země) z mezinárodního výzkumu ESS myslí, že lidé z jiných zemí, kteří přicházejí do země žít, většinou narušují kulturu země nebo ji obohacují, ale z jiného úhlu pohledu. Ukazuje, jak důležitý má proměnná přínos do vytvoření všech klastrů. Pokud bychom v tomto grafu identifikovali nesignifikantní přínos proměnné pro tvorbu všech výsledných shluků, měli bychom zvážit, zda má daná proměnná ještě dále pro analýzu mezinárodních dat nějaký význam a zda ji nevyloučit. Tímto způsobem může komparativní výzkum hledat signifikantní proměnné, které by na systémové úrovni mohly mít přínos pro vysvětlení mezinárodních odlišností.

## 9.4 Validita klastrování

Jeden z největších problémů, kterým výzkumníci ohledně klastrovacích technik v praxi čelí, je otázka stability a validity nalezených klastrů: *Dosáhneme stejné klasifikace i za použití jiných proměnných? Dosáhneme stejné klasifikace i v novém vzorku objektů? Skutečně odpovídali respondenti z různých klastrů odlišně ve stejných položkách?* Tato zjištění nejsou tolik potřebná pro obecný popis dat a explorační záměry, ale jsou nezbytná v případě, že nalezené skupiny mají zásadní význam pro výslednou konstrukci vědecké

typologie [Everitt 1979: 171]. Aldenderfer a Blashfield [1984] zařadili validizaci výsledného klasifikačního řešení do pěti základních kroků klastrovací analýzy a upozornili na to, že tento krok nelze nikdy opomíjet.

Garson [2010] navrhl k získání validních shluků, nehledě na jakýkoli zvolený klastrovací způsob, tři neformální pravidla:

- ◆ *velikost* – všechny klastry musí mít dostatečný počet objektů, aby byly smysluplné. Pokud výzkumník z analýzy získal velké množství malých klastrů, požadoval po algoritmu příliš mnoho shluků, naopak jeden velký dominantní klastř svědčí o výzkumníkově příliš nízkém požadavku na počet klastrů (popř. ukazuje na zvolení nevhodného měření vzdálenosti objektů).
- ◆ *smysluplnost* – ideálně, podobně jako ve faktorové analýze, by měl smysl každého klastru snadno a intuitivně vyplývat z jednotlivých proměnných, které jej vytváří
- ◆ *validita ke kritériu* – získané klastry by měly korelovat s teoretickým konceptem, který v podstatě klastrování reflektuje (vliv charakterizujících proměnných se tak očekává již z teoretických znalostí nebo předešlého výzkumu)

Vedle těchto neformálních obecných pravidel existuje pro ověřování platnosti shluků, výběr vhodné shlukovací metody a stanovení počtu shluků několik dalších již systematictějších přístupů. Mají podobu heuristických procedur nebo formálních testů (test absence struktury, test hierarchické struktury atp.), někdy jsou založeny na vizuálním posouzení grafů, jindy na indexech a statistikách.

Jako jeden z nejpřehlednějších systémů validizačních technik se zatím jeví třídění Aldenderfera a Blashfielda [1984: 62], kteří vymezili přibližně pět přístupů zpětné validizace. První dva jsou ryze statistické a podle jejich soudu ne zcela ideální, jedná se o *kofenetickou korelaci* a *multidimenzionální analýzu rozptylu* (MANOVA), které testují přesnost hierarchických stromů. Vedle těchto matematických analýz dále navrhuji známé *replikační řešení*, které využívá možnosti trénovacích a testovacích skupin. Problémem v tomto případě je, že opakující se klastrové řešení v různých vzorcích objektů ze stejné populace ještě nemusí znamenat důkaz validity klastrového řešení, neboť použité stejné klastrovací metody mohou opět vést k neúspěšnému řešení. Další alternativou je procedura, jež provádí *test signifikance* srovnáváním klastrů s jinými proměnnými (tzv. externí proměnné), než jaké vygenerovaly předchozí výsledky klasifikace. Poslední přístup, tzv. *Monte Carlo procedura*, využívá náhodného generování datového souboru se stejnými obecnými charakteristikami jako mají původní data a na oba datové soubory pak používá stejné klastrovací techniky, výsledná



řešení jsou porovnávána ANOVOU.

Důvodem řídkého používání validizačních technik v případě použití klastrové analýzy je ten, že metodologický design nutný ke sběru relevantních charakteristických dat pro validizaci je značně finančně náročný [Aldenderfer, Blashfield 1984: 66]. Další překážkou je samotná explorační povaha klastrové analýzy, která v zájmu svých vlastních průzkumných cílů a budování typologie tam, kde není v datech ani známá, ani patrná, nemůže být předem vylepšena. Nemůže svým způsobem předem promýšlet svou vlastní validizaci.

Je nutné si uvědomit, že výsledek shlukování je vždy ovlivněn výběrem proměnných a jejich vahami při analýze, případně jejich transformací. Svůj vliv mají také odlehlá pozorování a chybějící údaje, včetně primárních znalostí o populaci, které mohou nesprávně ovlivnit subjektivní výzkumníkův výběr počátečních klastrů. Klasifikace objektů nebo proměnných získaná shlukovou analýzou nemá žádnou vnitřní validitu. Oprávněnost každé klasifikace a z ní vycházející explanační struktury je zdůvodnitelná shodou se známými fakty bez vztahu ke způsobu jejího získání [Hendl 2004: 464].

Výsledky shlukovací metody jsou směsí struktury, kterou do dat vkládá sám algoritmus procedury, a té, jež je v datech skutečně přítomná. Používané procedury obsahují operace, které na jedné straně systematicky opomíjejí určité (pro každou proceduru individuální) rysy dat a na druhé straně protěžují ostatní. Při analýze se často opomíjejí dvě možnosti: data nemusí obsahovat žádné shluky (absence jakýchkoli diskriminačních proměnných, rovnoměrné rozložení bodů) anebo data mohou tvořit pouze jeden shluk (absence diskriminačních proměnných a vzájemná příbuznost údajů) [Hendl 2004: 465].

Množina získaných shluků není konečným výsledkem, ale pouze možným návrhem struktury, prozkoumáním velkého množství dat. Pro praktické cíle výzkumu tak ani nemusí být formální test signifikance řešení klastrů nezbytný, neboť je výhodnější, když se badatelé zaměřují na možnosti několika alternativních klasifikací, aby tak reflektovali různé aspekty dat a mohli je prohlubovat jinými technikami [Everitt 1979: 176].

## 10 Moderní klastrovací metody

Předešlé metody založené na měření vzdáleností objektů jsou příliš složité a náročné na výpočty. Navíc efektivně fungují maximálně do počtu 16 proměnných. Řešením problému s velkým počtem proměnných může být snížení rozměru úlohy (tzv. redukce dimenzí) provedená kupříkladu na základě analýzy hlavních komponent, která lineární kombinací původních proměnných vytvoří nové pomocné proměnné. Nebo pomocí vícerozměrného škálování zjistit skupinu podobných proměnných a z každé skupiny ponechat pouze jednu reprezentativní.

V současné době je čím dál větší důraz kladen na požadavky přiměřené náročnosti analytických technik, na schopnost pracovat s datovými soubory obsahujícími velký počet proměnných, navíc různých typů, a tyto požadavky jsou obzvláště důležité pro analýzu komparativních výzkumů. Současně s tím by se měly statistické shlukové techniky efektivně vyrovnávat s odlehlými pozorováními a kvalitně zjišťovat optimální počet shluků. Jejich výsledky by neměly být závislé na pořadí objektů v souboru a měly by být interpretovatelné, se schopností odhalovat překrývající se shluky [Řezanková, Húsek a Snášel 2009: 108]. Aktuálním trendem se tedy stává orientace na shlukování velkých databází a v rámci těchto moderních přístupů probíhá modifikace všech známých klasifikačních technik, které byly do této doby aplikovány. Pokusme se o některých z nich stručně zmínit.

Podkladem pro *modifikaci metod rozpadu* se stávají hybridní klasifikace, které přiřazují objekty z určité podmnožiny do  $k$  shluků, po němž následuje zařazení ostatních objektů do některého ze vznikajících shluků a to podle určitého optimalizačního kritéria. Nejznámější jsou metody **CLARA** (Clustering **LAR**ge Application) a **CLARANS** (Clustering Large Application based on **RAN**domised Search), vycházející z metody shlukování kolem medoidů PAM a provádějící proceduru pouze na reprezentativní části vzorku, obvykle náhodně vybraném. Vyhledané medoidy se pak používají pro kompletaci shluků celé databáze.

Metoda **PAM** (Partitioning Around Medoids) by ve srovnání s metodou K-průměrů měla být více robustní, medoidy jako zástupci shluků nejsou totiž tolik citlivé k výskytu extrémních odlehlých hodnot, které vstupují do výpočtu průměru. Počáteční medoid je určen tak, aby součet vzdáleností jednotlivých objektů ve shluku od tohoto vybraného objektu byl minimální, objekty se postupně přiřazují k nejbližším medoidům. Střed shluku určený jako (nejspíše fiktivní) objekt s průměrnými hodnotami jednotlivých sledovaných proměnných přes objekty zařazené v daném shluku může být poměrně daleko od skutečných objektů.

Navíc je výhodou PAM, že není potřeba na úvod zadávat nebo náhodně vybírat množinu reprezentativních případů, jsou totiž v procesu výpočtu PAM poměrně inteligentně nalezeny [Řezanková, Húsek a Snášel 2009: 87].

Mezi množstvím alternativ ke klastrové analýze patří tzv. expected maximization clustering (**EM**). Tato metoda je značně podobná K-means klastrování až na to, že používá maximální pravděpodobnostní odhad (maximum likelihood estimation) k determinaci shluků. Objekty jsou přiřazovány ke klastrům na základě maximální pravděpodobnosti distribuce dat, produkující finální klastr. Díky využití maximální věrohodnostní funkce může EM podporovat jak kategoriální, tak i spojitě proměnné, což je jeden z důvodů jejího použití [Garson 2010].

*Modifikace hierarchických metod* má nejčastěji podobu tzv. frakcionalizace. Jedná se o rozdělení datového souboru do podsouborů (frakcí) a aplikování shlukové analýzy na každou tuto frakci. Pro každý shluk se spočte centroid, shluky ve frakcích jsou dále shlukovány do  $k$  skupin stejnou metodou shlukové analýzy (počet  $k$  musí být stanoven předem), zbylé objekty se přiřazují do vytvořených  $k$  shluků na základě centroidů. Jistou modifikovanou podobu hierarchických metod má již dvoukroková shlukovací analýza, protože je založena na algoritmu **BIRCH**, který uspořádává objekty do podshluků, jež jsou charakterizovány pomocí shlukovacích vlastností CF a tyto vlastnosti jsou dále shlukovány do  $k$  skupin pomocí hierarchické shlukové analýzy.

Mezi další moderní metody založené na modifikaci hierarchických technik se řadí metoda **CURE** (Clustering Using **R**epresentatives) pro spojitě proměnné a metoda **ROCK** (**R**Obust Clustering using lin**K**s) pro kategoriální proměnné. Metoda CURE vybírá z databáze vzorek dat a rozčleňuje ho do zvolených částí. Na základě parciálního shlukování v každé části technika nalézá reprezentativní body pro každý shluk a v následném kroku provádí kompletní shlukování v celém vzorku podle vzdálenosti reprezentačního bodu od shluku. Eliminuje jakékoli odchylky shlukovacího procesu. Roste-li shluk příliš pomalu, odstraní se v průběhu shlukování, navíc i malé shluky jsou na konci shlukovací fáze eliminovány [Řezanková, Húsek a Snášel 2009: 114]. Metoda ROCK spočívá v náhodném výběru objektů, které se shlukují do požadovaného počtu shluků, po němž následuje přiřazení zbylých objektů. ROCK využívá stejné principy jako shlukování nejbližších sousedů, základními používanými prostředky jsou sousedi (na základě podobnosti) a vazby (links) jako počet společných sousedů těchto bodů, jejichž součet má za úkol metoda ROCK maximalizovat [Řezanková, Húsek a Snášel 2009: 119].

Současnými moderními technikami se také staly *metody založené na hustotě, mřížce a modelu*. Vycházejí ze zobrazení objektů jako bodů v m-rozměrném prostoru. Shluky jsou definovány jako oblasti ve výběrovém prostoru, které se vyznačují větší hustotou bodů ve srovnání s ostatními oblastmi. Vychází se z předpokladu, že každý objekt má ve výběrovém prostoru své vlastní sousedství, ze kterého je přímo dosažitelný. Pro objekt ve shluku platí, že jeho sousedství (dané poloměrem) musí obsahovat alespoň minimální počet dalších objektů a pak se nazývá jádro. Metody tohoto typu mohou být použity pro zjišťování šumu (nepatří do žádného shluku) a odlehlých objektů a k odhalení shluků libovolného tvaru [Řezanková, Húsek a Snášel 2009: 114].

Posledním příkladem moderního vývoje je *shlukování podprostorů*. Shlukování podprostorů je vhodné použít pro velké množství objektů a proměnných, neboť namísto vytváření nových proměnných zkoumá podprostory původního prostoru. Základem tohoto procesu jsou metody založené na hustotě. Cílem je nalézt podmnožiny proměnných tak, aby soubory objektů zahrnovaly oblasti s vysokou hustotou. Tím pádem vznikají shluky jako oblasti navazujících jednotek s vysokou hustotou v rámci určitého podprostoru [Řezanková, Húsek a Snášel 2009: 118].

Velké množství nově se rozvíjejících shlukovacích metod je ještě stále softwarově těžko dostupné a také uživatelsky velice náročné, obzvláště kvůli své statistické sofistikovanosti. Kupříkladu nejmodernější biologicky inspirované metody shlukování jako umělé neuronové sítě (Kohenovy mapy atp.), algoritmy rostoucích buněk, genetické algoritmy, které byly vytvořeny jako alternativa k čistě matematickým shlukovacím metodám, jsou dostupné jen v rámci specializovaných počítačových programů [více Řezanková, Húsek, Snášel 2009].

Analýza zmiňovanými moderními algoritmy částečně dostupnými například v programu R, které zatím v komerčních softwarových nástrojích nejsou tolik rozšířené, řeší některé problémy klasických shlukovacích metod. Společně s tím také nabízí pro odhad užitečnosti modelu při jeho případné aplikaci (včetně možnosti výběru nejvhodnějšího modelu) zajímavý a poměrně málo rozšířený výstup, totiž graf obrysů (Silhouette plot). Nabízené klady, které moderní přístupy přináší, jsou neoddiskutovatelné, a také se stávají inspirací pro další sociálněvědecké bádání v této oblasti, obzvláště v jejich efektivním užití na rozsáhlé databáze komparativních výzkumů. Avšak pro mnohé praktické aplikace jsou méně sofistikované a méně technicky pokročilé metody postačující.

## Závěr

Komparativní výzkumy specifikují vědci jako *sociálněvědní bádání*, které studuje stejné nebo *ekvivalentní sociální jevy* ve dvou či více společnostech z důvodu užití společnosti (resp. kultury, země, národu, systému, státu) jako jedné z nezávislých proměnných v *analýze mezinárodních dat*. Tato studie se soustředila na tři širší rámce, v nichž jsou komparativní výzkumy považovány za skutečný přínos pro teoretickou i empirickou oblast sociologie.

Prvním nejširším rámcem, který zároveň zastřešuje i zbylé dva celky, je teoretický rámec. Podstatu teoretického okruhu vystihuje sousloví *sociálněvědní bádání* z výše uvedeného pojetí komparativních výzkumů. V této studii bylo předvedeno, jaký neocenitelný význam mají komparativní výzkumy jako sociálněvědní bádání pro vytváření obecné sociální teorie a pro její empirickou ověřitelnost, jež jsou schopny testovat realizaci výzkumů v různých sociálních a kulturních prostředích a následným porovnáváním získaných výsledků. Bylo prezentováno, že jejich primárním teoretickým cílem je namísto pouhého kategorizování rozdílů mezi sociálními systémy hledat unikátní společné aspekty a vlastnosti, které by zkoumané jevy vymanily z jejich sociokulturního prostředí a umožnily formovat obecné teorie sociálních fenoménů a procesů ve společnosti. Získáváním hlubších znalostí o sociální realitě se tak výzkumníkům zprostředkovává i lepší porozumění odlišnostem jednotlivých společností, jejich struktuře a kulturním zvykům.

Tato práce poukázala na to, že povaha sociálněvědního (mezinárodního) bádání se značně liší podle výběru objektu výzkumu, geografického rozsahu výzkumu, časového vymezení a že jeho charakter ovlivňuje i úroveň, na které jsou pozorování z různých kultur a zemí analyzována. Bylo upozorněno na skutečnost, že tato sociálněvědní bádání mají svou specifickou historii, ve které se z pouhého srovnávání výstupů národních šetření, realizovaných v různých sociálních podmínkách a navíc odlišnými výzkumnými strategiemi, vyvinuly rozsáhlé a promyšlené mezinárodní výzkumné programy, které systematicky a kontinuálně provádí sociální šetření v mnoha zemích světa. Teoretický přínos komparativních výzkumů byl také zachycen srovnáním logiky národních a mezinárodních šetření, ze kterého vyplývá, že komparativní výzkumy jsou hodnotné především svou schopností přispívat k budování obecných teorií a akumulovat různorodé poznatky o sociálním světě. Zároveň jsou také během posuzování odlišných sociálních a kulturních procesů výzkumníci nuceni promýšlet svou vlastní národní kulturu z perspektivy vnějšího pozorovatele, což je bohatým zdrojem teoretických poznatků v oblasti národní vědy.

Druhým užším rámcem studie byl metodologický okruh. Metodologická oblast komparativních výzkumů je primárně koncipována tak, aby byla schopna všech teoretických cílů mezinárodních šetření efektivně dosáhnout. V centru pozornosti této studie stálo předvedení komparativních výzkumů jako sociálních výzkumů, které uplatňují kvantitativní metody jak na poli primárního mezinárodního výzkumu, tak i v oblasti sekundárních analýz mezinárodních dat. Namísto výzkumu makrosociální povahy, který se pojí s řadou metodologických problémů, upřednostňují komparativní výzkumy volbu teorie středního dosahu a vybírají si pro své studium sociální jevy ze *dvou a více společností*, což značně ulehčuje i následnou interpretaci jejich výsledků.

Z metodologické perspektivy srovnání logiky národních a mezinárodních šetření bylo poukázáno na to, že charakter vědních disciplín by měl být flexibilní a přijmout spolupráci těchto dvou oblastí sociálního výzkumu, zvláště ve sféře dvousměrného transferu výzkumných znalostí a standardů. Toto vzájemné akceptování metodologických pravidel ale musí být řádně zváženo, aby nedocházelo obzvláště při realizaci mezinárodního šetření k narušení *ekvivalentního srovnávání* sociálních jevů. Otázce ekvivalence výzkumu byla věnována speciální pozornost, aby se tak zdůraznila její důležitost v celé metodologické koncepci realizace komparativních výzkumů. Požadavky mezinárodních šetření na identifikaci srovnatelných jevů a ekvivalentní měření vybraných teoretických konceptů ve všech zkoumaných společnostech jsou neodmyslitelnými prvky všech metodologických otázek, jež komparativní výzkumy provázejí. Tato studie se snažila primárně doložit závažnost propojení reliability a validity výzkumu s jeho ekvivalencí v mezinárodním kontextu.

Ve studii bylo také demonstrováno, jak důležité je věnovat zvýšenou pozornost dosahování kvality v mezinárodních výzkumech, obzvláště ve vztahu k interpretačně hodnotným výstupům analýzy mezinárodních dat. Kvality lze v komparativních výzkumech dosáhnout synchronizovaných uplatněním třech specifických přístupů (statistického, technicko-administrativního a psychometrického), avšak hrozí zde také, že každá země upřednostní jen některý z těchto zdrojů kvality výzkumu a nebo pouze jeden jeho specifický rys. Realizace ekvivalentního a kvalitního mezinárodního šetření se tak potýká se zakořeněnými výzkumnými zvyklostmi zemí, jež na něm participují. Proti specifickým výjimkám, které by se ve výzkumu objevovaly, kdyby se participujícím zemím dovolila přílišná flexibilita ve volbě výzkumových strategiích, nasazují mezinárodní výzkumné programy přísné metodologické standardy svých šetření. Ty se pak musejí plošně dodržovat ve všech

zemích, aby se případné nesrovnatelnosti eliminovaly. Jak náročnými postupy se v současných mezinárodních šetřeních zabezpečuje kvalita názorně dokumentuje promyšlený design mezinárodního sociálního výzkumného programu European Social Survey prezentovaný v příslušné části práce.

Třetím a nejuzšíším rámcem této práce, v němž komparativní výzkumy také představují skutečný teoretický a empirický přínos do sociologické vědní disciplíny, byl empirický proces *analýzy mezinárodních dat*. I základní povaha a výchozí cíle srovnávacích analýz jasně odpovídají teoretickým ambicím komparativních výzkumů, neboť jejich prvotní záměr se vyznačuje soustavným hledáním vlivu systémových faktorů na individuální chování jedinců v různých kulturách. Analýzy komparativních výzkumů se primárně nezaměřují na pouhé kategorizování sociálních systémů a sociálních jevů, ale jejich cílem je nalézt vysvětlení pro toto seskupení podobných a odlišných zemí a zastoupit názvy zemí nalezenými vysvětlujícími proměnnými. Z tohoto důvodu se v analýze mezinárodních dat užívá společnost (resp. kultura, země, národ, systém, stát) jako jedna z nezávislých proměnných.

Tato studie poukázala na to, že pro dosažení vytyčených cílů komparativních šetření je při statistické analýze velkého množství dat důležité využít možností kombinace multidimenzionálních technik. Optimálním řešením je přitom aplikace vhodných exploračních technik (např. shlukové analýzy) k vytvoření prvního přehledu o velkém množství dat a následné rozšíření a prohloubení získaných poznatků z výsledků analýz užitím dalších multidimenzionálních technik, které se spíše zaměřují na podrobnější analýzu vztahů zvolených proměnných (např. faktorová analýza, regresní analýza). V této práci byl důraz položen na jeden z prvních kroků analýzy, při němž je vhodné využít explorační potenciál shlukové analýzy. Ta jako jedna z klasifikačních metod dokáže odhalit skrytou strukturu v datech a odkrýt jejich podstatné parametry. To se může následně stát dobrým startovacím bodem (resp. teoretickým základem) pro pokročilejší typy analýz, které povedou k nalezení vysvětlujících proměnných na systémové úrovni a možnostem vytvoření obecné sociální teorie.

Konkrétně byly v této souvislosti předvedeny analytické schopnosti tří vybraných klastrovacích technik (hierarchické aglomerativní metody, metody K-průměrů a dvoukrokové shlukovací techniky) na konkrétních mezinárodních datech první vlny European Social Survey tak, aby podtrhly užitečnost prvotního průzkumu v mezinárodních datech a naznačily eventuelní využití získaných informací v hlubších statistických analýzách, a ještě více se tak zdůraznil skutečný přínos empirické analýzy mezinárodních dat do teoretické oblasti sociální vědy, kterému je důležité se i nadále systematicky věnovat.

Věřím, že tato práce přispěla k vytvoření komplexnějšího pohledu na problematiku komparativních výzkumů, neboť v takto souhrnném pojetí je sociální vědy v České republice příliš systematicky nereflektují. Vzhledem k současnému stavu metodologické rozpracovanosti mezinárodních výzkumů lze očekávat rostoucí zájem o jejich teoretický a empirický přínos v oblasti sociálních věd. Jsem přesvědčena, že nastíněné výzvy komparativních mezinárodních šetření pro sociální výzkum mohou iniciovat kontinuální zapojení České republiky do takovýchto druhů empirických šetření.



## Seznam použité literatury

- Aldenderfer, Mark S., Roger K. Blashfield. 1984. *Cluster Analysis*. Newbury Park, CA: Sage.
- Almond, Gabriel A., Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton, NJ: Princeton University Press.
- Bailey, Kenneth D. 1994. *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, CA: Sage.
- Banfield, J. D., Adrian E. Raftery. 1998. „Modl-based Gaussian and non-Gaussian clustering.“ *Biometrics* 49: 803-821.
- Barker, D. 1976. „Hierarchic and Non-Hierarchic Grouping Methods: An Empirical Comparison of two Techniques.“ *Geografiska Annaler. Series B, Human Geography* [online] 58 (1): 42 – 58 [04/11/2009]. Dostupné z: <<http://www.jstor.org/stable/490777>>.
- Billiet, Jaak. 2003. „Cross-cultural equivalence with structural equation modeling.“ Pp. 247 – 263 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Billiet, Jaak, Jerry Welkenhuysen-Gybels. 2003. „Assesing cross-national construct equivalence in the ESS: The case of six immigration items.“ *Journal of Cross-Cultural Psychology* 34 (6): 702-722.
- Bollen, Kenneth A., Barbara Entwisle, Arthur S. Alderson. 1993. „Macrocomparative Research Methods“. *Annual Review of Sociology* [online] 19: 321 – 351 [05/04/2010]. Dostupné z: <<http://www.jstore.org/stable/2083391>>.
- Borg, Ingwer. 1998. „A Facet-Theoretical Approach to Item Equivalency.“ *ZUMA – Nachrichten Spezial* 3, January 1998: 145 – 158.
- Braun, Michael. 2003a. „Communication and social cognition.“ Pp. 57 – 67 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Braun, Michael. 2003b. „Errors in comparative survey research: An overview.“ Pp. 137 – 142 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Campbell, D.T., D. W. Fiske. 1959. „Convergent and Discriminant Validation by the Multimethod-Multitrait Matrix.“ *Psychological Bulletin* 56: 833 – 853.
- Cartwright, Nancy. 2002. „The Limits of Causal Order, From Economics to Physics.“ Pp. 137 - 51 in U. Mäki (ed.). *Fact and Fiction in Economics: Models, Realism, and Social Construction*. Cambridge: Cambridge University Press.

- Couper, Mick P., Edith D. de Leeuw. 2003. „Nonresponse in cross-cultural and cross-national survey.“ Pp. 157 – 177 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Dansecu, Evangeline R., Wayne Holden. 1998. „Are There Different Types of Homeless Families? A Typology of Homeless Families Based On Cluster Analysis.“ *Family Relations* [online] 47 (2): 159 - 165 [08/03/2010]. Dostupné z: <<http://www.jstor.org/stable/585620>>.
- De Leeuw, Edith D., Joop J. Hox. 2003. „The Use of Meta-analysis in Cross-national Studies.“ Pp. 329 – 345 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Dillman D. A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- DiStefano, Christine, R. K. Kamphaus. 2006. „Investigating Subtypes of Child Development: A Comparison of Cluster Analysis and Latent Class Cluster Analysis in Typology Creation.“ *Educational and Psychological Measurement* [online] 66 (5): 778 – 794 [29/10/2009]. Dostupné z: <<http://epm.sagepub.com/cgi/content/refs/66/5/778>>.
- Driver, H. E., A. L. Kroeber. 1932. „Quantitative expression of culture relationships.“ *University of California Publications in American Archeology and Ethnology* 31: 211 – 256.
- Edelbrock, C. 1979. „Comparing the accuracy of hierarchical clustering algorithms: the problem of classifying everybody.“ *Multivariate Behavioral Research* 14: 367 – 384.
- ESS. 2002a. ESS Working Paper. *Chapter 1: European Social Survey Core Questionnaire Development* [online]. Dostupné z: <<http://www.europeansocialsurvey.org/>> (The Questionnaire, Core Questionnaire Development) [03/10/2009].
- ESS. 2002b. European Social Survey, Round 1: *Documentation Report Edition 6.0* [online]. The ESS Data Archiv. Dostupné z: <<http://ess.nsd.uib.no/ess/round1/>> (ESS Round 1 – 2002, Survey documentation) [21/03/2010].
- ESS. 2003. European Social Survey, Round 2: *Sampling for the European Social Survey – Round II: Principles and requirements* [online]. Dostupné z: <<http://www.europeansocialsurvey.org/>> (Methodology, Sampling Designs - Round 2) [03/10/2009].
- ESS. 2004. European Social Survey, Round 1: *Technical Report Edition 2. Chapter 2. The Sample* [online]. Dostupné z: <<http://www.europeansocialsurvey.org/>> (Data Documentation, Technical Report) [21/03/2010].
- ESS. 2007. European Social Survey, Round 4: *Specification for participating countries* [online]. Dostupné z: <<http://www.europeansocialsurvey.org/>> (Project Specification - Round 4) [16/04/2009].

- Everitt, B. S. 1979. „Unresolved Problems in Cluster Analysis.“ *Biometrics* [online] 35 (1): 169 – 181 [08/03/2010]. Dostupné z: <<http://www.jstor.org/stable/2529943>>.
- Fitzgerald, Rory, Sally Widdop, Michelle Grey, Debbie Collins. 2009. „Testing for equivalence using cross-national cognitive interviewing.“ *Working Paper Series 1* [online], European Social Survey. London: Center for Comparative Social Surveys. Dostupné z: <<http://www.europeansocialsurvey.org/>> (Cognitive Pretesting) [03/10/2009].
- Fraley, Chris, Adrian E. Raftery. 2002. „Model-Based Clustering, Discriminant Analysis, and Density Estimation.“ *Journal of the American Statistical Association* [online] 97 (458): 611 – 631 [08/03/2010]. Dostupné z: <<http://www.jstore.org/stable/3085676>>.
- Ganninger, Matthias. 2006. *Estimation of Design Effects for ESS Round 2*. Documentation [online], European Social Survey. Dostupné z: <http://www.europeansocialsurvey.org/> (Methodology, Sampling Designs - Round 2) [03/10/2009].
- Garson, David G. 2010. „Cluster Analysis.“ *Statnotes: Topics in Multivariate Analysis* [online]. Dostupné z: <<http://faculty.chass.ncsu.edu/garson/pa765/statnote.html>> [08/03/2010].
- Goldthorpe, John H. 1997. „Current Issues in Comparative Macrosociology: A Debate on Methodological Issues“. *Comparative Social Research* 16: 1 – 26.
- Gower, J. C. 1967. „A Comparison of Some Methods of Cluster Analysis.“ *Biometric* [online] 23 (4): 623 - 637 [08/03/2010]. Dostupné z: <<http://jstor.org/stable/2528417>>.
- Groves, R. M. 2004. *Survey Errors and Survey Costs*. New York: Wiley.
- Hantrais, Linda, Steen Mangen. 2006. *Cross-National Research Methods in the Social Sciences*. London/New York: Pinter.
- Harkness, Janet. 2003. „Questionnaire translation.“ Pp. 35 – 56 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Harkness, Janet, Fons J. R. Van de Vijver, Timothy P. Johnson. 2003. „Questionnaire design in comparative research.“ Pp. 19 – 34 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Harkness, Janet A., Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Harkness, Janet, Peter Ph. Mohler, Fons J. R. Van de Vijver. 2003. „Comparative Research.“ Pp. 3 – 16 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.

- Häder, Sabine, Siegfried Gabler. 2003. „Sampling and estimation.“ Pp. 117 – 134 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Hendl, Jan. 2004. *Přehled statistických metod zpracování dat. Analýza a metaanalýza dat*. Praha: Portál.
- Herkenrath, Mark. 2002. „Quantitative Cross-National Analysis as a Research Tool in the Sociology of Developing Countries: A Critical Examination.“ *Current Sociology* [online] 50 (4): 517 – 530 [24/10/2009]. Dostupné z: <<http://csi.sagepub.com/cgi/content/abstract/50/4/517>>.
- Hopkins, Terence K., Immanuel Wallerstein. 1967. „The Comparative Study of National Societies“. *Social Science Information* 5 (1): 25 – 58.
- Hyman, H. H. 1972. *Secondary Analysis of Sample Surveys: Principles, Procedures, and Potentialities*. New York: Wiley in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Christiansen, Bjørn, Herbert Hyman, Ragnan Rommetveit. 1951. *Cross-National Social Research*. Oslo: Institut for Social Research.
- Inkeles, Alex, Peter Rossi. 1956. „National Comparison of Occupation Prestige.“ *American Journal of Sociology* 61 in Adam Przeworski, Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Jacobson, E., H. Kumata, J. E. Gullahorn. 1960. „Cross-cultural Contribution to Attitude.“ *Research. Public Opinion Quarterly* 24: 205 – 223 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Johnson, Timothy P. 1998. „Approaches to Equivalence in Cross-Cultural and Cross-National Survey Research.“ *ZUMA – Nachrichten Spezial* 3, January 1998: 1 – 40.
- Johnson, Timothy P., Fons J. R. Van de Vijver. 2003. „Social desirability in cross-cultural research.“ Pp. 195 – 204 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Jowell, Roger. 1998. „How comparative is comparative research?“ Centre For Research Into Elections And Social Trends *CREST. Paper No 66* [online]. *Maison Francaise*, Oxford: June 26/27 1997 [27/10/2009]. Dostupné z: <<http://www.crest.ox.ac.uk/papers.htm>>.
- Jowell, Roger. *About the European Social Survey*. ESS Brochure [online]. London: Limehouse [03/10/2009]. Dostupné z: < <http://www.europeansocialsurvey.org/> > [03/10/2009].

- Jowell, Roger and the Central Co-ordinating Team. European Social Survey 2002/2003: Technical Report [datový soubor]. London: Centre for Comparative Social Surveys, City University (2003). Dostupné z: <<http://ess.nsd.uib.no/ess/round1/download.html>> (Integrated file, edition 6.1) [21/03/2010].
- Kish, Leslie. 1994. „Multipopulation Survey Designs: Five Types with Seven Shared Aspects.“ *International Statistics Review* 62: 167 - 186 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). 2003. Cross-cultural survey methods. New Jersey: John Wiley & Sons.
- Kittel, Bernhard. 2006. „A Crazy Methodology?: On the Limits of Macro-Quantitative Social Science Research.“ *Internacional Sociology* [online] 21 (5): 647 – 677 [29/10/2009]. Dostupné z: <<http://iss.sagepub.com/cgi/content/refs/21/5/647>>.
- Kohn, Melvin L. 1989. *Cross – National Research in Sociology*. Newbury Park, CA: Sage.
- Krejčí, Jindřich. 2006. „Mezinárodní sociálněvědní komparativní výzkum a Česká republika: Přehled výzkumů a dostupných dat.“ *Sociologický časopis* 42 (1): 149 – 173.
- Krejčí, Jindřich. 2008. *Kvalita sociálněvědních výběrových šetření v České republice*. Praha: Sociologické nakladatelství.
- Landau, Sabine, Brian S. Everitt. 2004. *A Handbook of Statistical Analysis using SPSS*. Boca Raton: Chapman & Hall/CRC Press LLC.
- Lievesley, Denise. 2001. Making a Difference: A Role for the Responsible International Statistician? *Journal of the Royal Statistical Society. Series D (The Statistician)* [online] 50 (4): 367 – 406 [29/10/2009]. Dostupné z: <<http://www.jstore.org/stable/2681223>>.
- MacQueen, J. 1967. „Some methods for classification and analysis of multivariate observation.“ Pp. 281 – 297 in L. M. LeCam, J. Neyman (eds.). *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1*. Berkeley: University of California Press.
- Madden, Ch. 2004. „Making Cross-country Comparisons of Culture Statistics: Problems and Solutions (International Comparisons of Culture Sectors: An Exploratory Investigation).“ *IFACCA, Australian Council for the Arts* [online]. Sydney, 2004 [04/10/2009]. Dostupné z: <[www.australiacouncil.gov.au/](http://www.australiacouncil.gov.au/)>.
- Marsh, Robert M. 1967. *Comparative Sociology: A Codification of Cross-Societal Analysis*. New York: Harcourt, Brace & World.
- Maxwell, Bruce A., Frederic L. Pryor, Casey Smith. 2002. „Cluster Analysis in Cross-cultural Research.“ *World Cultures* 13 (1): 22 – 38.
- McQuitty, L. L. 1957. „Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies.“ *Educational and Psychological Measurement* 17: 207 – 229.

- Melia, M., D. Heckerman. 1998. *An experimental comparison of several clustering and initialization methods*. Microsoft Research Technical Report MSR-TR-98-06.
- Merton, R. K. 2000. *Studie ze sociologické teorie*. Praha: Sociologické nakladatelství.
- Mills, Melinda, Gerhard G. Van de Bunt, Jeanne de Bruijn. 2006. „Comparative Research: Persistent Problems and Promising Solutions.“ *International Sociology* [online] 21 (5): 619 – 631 [24/10/2009]. Dostupné z: <<http://iss.sagepub.com/cgi/content/refs/21/5/619>>.
- Mohler, Peter Ph., Rolf Uher. 2003. „Documenting comparative surveys for secondary analysis.“ Pp. 311 – 327 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Morris-Jones, W. H. 1964. *The Government and Politics of India*. New York: Doubleday in Adam Przeworski, Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Motoyama, Yasuyuki. 2008. „What Was New About the Cluster Theory?: What Could It Answer and What Could It Not Answer?“ *Economic Development Quarterly* [online] 22 (4): 353 – 363 [29/10/2009]. Dostupné z: <<http://edq.sagepub.com/cgi/content/abstract/22/4/353>>.
- Naroll, Raoul. 1965. „Galton's Problem: The Logic of Cross-Cultural Analysis.“ *Social Research* 32 in Adam Przeworski, Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Nowak, Stefan. 1989. „Comparative Studies and Social Theory (The gap between the amount of broad sociological data and the present state of social theory).“ Pp. 34 – 56 in Melvin L. Kohn (ed.). *Cross-national Research in Sociology*. Newbury Park, CA: Sage.
- Plecitá-Vlachová, Klára, František Kalvas. 2005. „The European Social Survey (ESS): o projektu a dostupnosti dat.“ *SDA Info* VII (2): 14 – 16.
- Przeworski, Adam, Henry Teune. 1966. „Equivalence in Cross-National Research.“ *The Public Opinion Quarterly* [online] 30 (4): 551 – 568 [24/10/2009]. Dostupné z: <<http://jstor.org/stable/2476962>>.
- Przeworski, Adam, Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Ragin, Charles. 1989. „New Directions in Comparative Research.“ Pp. 57 – 76 in Melvin L. Kohn (ed.). *Cross-national Research in Sociology*. Newbury Park, CA: Sage.
- Rencher, Alvin C. 2002. *Methods of multivariate analysis. Second edition*. New York: Wiley.
- Robinson, W. S. 1950. „Ecological correlation and the Behavior of Individuals.“ *American Sociological Review* 15: 351 – 357.

- Rokkan, Stein. 1964. „Reviewed work. The Civic Culture: Political Attitudes and Democracy in Five Nations by Gabriel A. Almond; Sidney Verba.“ *The American Political Science Review* [online] 58 (3): 676 – 679 [16/03/2010]. Dostupné z: <<http://www.jstor.org/stable/1953147>>.
- Rokkan, S., S. Verba, J. Viet, E. Almasy (eds.). 1969. *Comparative Survey Analysis*. Paris: Mouton in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Ross, C. E., J. Mirowski. 1984. „Socially-Desirable Response and Acquiescence in a Cross-Cultural Survey of Mental Health.“ *Journal of Health and Social Behavior* 25: 189 – 197 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Řezanková, Hana, Dušan Húsek, Václav Snášel. 2009. *Shluková analýza dat*. Praha: Professional Publishing.
- Saris, Willem E. 2003a. „Multitrait-Multimethod studies.“ Pp. 265 – 274 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Saris, Willem E. 2003b. „Response function equality.“ Pp. 275 – 288 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Singer, Milton. 1961. „A Survey of Culture and Personality Theory and Research.“ in Bert Kaplan (ed.). *Studying Personality Cross-Culturally*. Row, Peterson, Evanston, Ill. in Adam Przeworski, Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Skjåk, Knut Kalgraff, Janet Harkness. 2003. „Data collection methods.“ Pp. 179 – 193 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Smith, Tom W. 2003. „Developing Comparable questions in Cross-national survey.“ Pp. 69 – 91 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Sneath, P. H. A., R. R. Sokal. 1973. *Numerical Taxonomy*. San Francisco: W. H. Freeman.
- Straus, Murray A. 1968. „Society as a Variable in Comparative Study of the Family by Replication and Secondary Analysis.“ *Journal of Marriage and Family* [online] 30 (4): 565 – 570 [08/03/2010]. Dostupné z: <<http://www.jstore.org/stable/349495>>.
- Szakolczai, Arpad, Laszlo Fustos. 1998. „Value Systems in Axial Moments: A Comparative Analysis of 24 European Countries.“ *European Sociological Review* [online] 14 (3): 211 – 229 [08/03/2010]. Dostupné z: <<http://www.jstor.org/stable/522681>>.

- Šarmanová, Jana. 2002. *Metody dolování znalostí z dat*. Skrypta pro Katedru informatiky FEI VŠB-TU Ostrava. Brno: Datakon.
- Tilly, Charles. 1984. *Big Structure, Large Processes, Huge Comparisons*. New York: Russell Sage.
- Tourangeau R., L. J. Rips, K. A. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Tryon, R. C. 1939. *Cluster Analysis: Correlation Profile and Orthometric (Factor) Analysis for the Isolation of Unities in Mind and Personality*. Ann Arbor, MI: Edwards Brothers.
- Van de Vijver, Fons J. R., K. Leung. 1997. *Methods and Data Analysis for Cross-Cultural Research*. Newbury Park, CA: Sage in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler. 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Van de Vijver, Fons J.R. 2003a. „Bias and equivalence: Cross-cultural perspectives.“ Pp. 143 – 155 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Van de Vijver, Fons J.R. 2003b. „Bias and substantive analysis.“ Pp. 207 – 233 in Janet A. Harkness, Fons J. R. Van de Vijver, Peter Ph. Mohler (eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Vinopal, Jiří. 2008. *Kognitivní přístupy v metodologii výzkumných šetření: metoda okamžité validizace*. Praha: Sociologický ústav AV ČR.
- Wolfe, David A., Meric S. Gertler. 2004. „Cluster from the Inside and Out: Local dynamics and Global Linkages.“ *Urban Studies* [online] 41 (5 – 6): 1071 – 1093 [08/03/2010]. Dostupné z: <<http://usj.sagepub.com/cgi/content/refs/41/5-6/1071>>.
- Wolfson, Murray, Zagros Madjd-Sadjadi, Patrick James. 2004. „Identifying National Types: A Cluster Analysis of Politics, Economics, and Conflict.“ *Journal of Peace Research* [online] 41 (5): 607 – 623 [08/03/2010]. Dostupné z: <<http://www.jstor.org/stable/4149616>>.
- Worrell, Frank C., Beverly J. Vandiver, Barbara A. Schaefer, William E. Cross, Jr., Peony E. Fhagen-Smith. 2006. „Generalizing Nigrescence Profiles: Cluster Analyses of Cross Racial Identity Scale (CRIS) Scores in Three Independent Samples.“ *The Counseling Psychologist* [online] 34 (4): 519 - 547 [30/10/2009]. Dostupné z: <<http://tcp.sagepub.com/cgi/content/refs/34/4/519>>.
- Zubin, J. A. 1938. „A technique for measuring likemindedness.“ *Journal of Abnormal and Social Psychology* 3: 508 – 516.
- Žák, Libor. 2004. „Shluková analýza I.“ *Automatizace* 47 (3): 180 – 190.



## Seznam tabulek a grafů

### Seznam tabulek

Tabulka 4.1	Metody získávání ekvivalence v mezinárodních výzkumech
Tabulka 9.1	Aglomerační tabulka (Wardova metoda)
Tabulka 9.2	Příslušnost objektů ke klastrům (Cluster membership)
Tabulka 9.3	Tabulka ANOVY (test signifikance proměnných)
Tabulka 9.4	Final Cluster Centers (průměry proměnných v klastrech)
Tabulka 9.5	Cluster membership (vzdálenosti objektů od shluku)
Tabulka 9.6	Tabulka Cluster Distribution (příklad odlehlého klastru)
Tabulka 9.7	Tabulka Auto-Clustering: informační kritérium BIC

### Seznam grafů

Graf 9.1	Dendrogram: Euklidovská vzdálenost
Graf 9.2	Dendrogram: Minkowského vzdálenost
Graf 9.3	Metoda nejbližšího souseda (single linkage)
Graf 9.4	Metoda nejvzdálenějšího souseda (complete linkage)
Graf 9.5	Metoda průměrné vazby (Between-groups linkage)
Graf 9.6	Metoda průměrné vazby (Within-groups linkage)
Graf 9.7	Centroidní metoda
Graf 9.8	Mediánová metoda
Graf 9.9	Příklad jevu crossover u ne-monotonických metod
Graf 9.10	Wardova metoda

Graf 9.11	Příklad dendrogramu
Graf 9.12	Příklad vertikálního rampouchového grafu
Graf 9.13	Scatter Plot matrix (dvě proměnné) – Wardova metoda
Graf 9.14	Scatter Plot (dvě proměnné) – Wardova metoda
Graf 9.15	Scatter Plot K-průměrů (dvě proměnné)
Graf 9.16	Scatter Plot K-průměrů (vzdálenosti objektů od centra shluku)
Graf 9.17	Sloupcový diagram relativních četností kategorií pro kategorickou proměnnou
Graf 9.18	Sloupcový diagram vnitroklastrového rozptylu pro spojitě proměnné
Graf 9.19	Důležitost přínosu kategorických proměnných do klastru 1
Graf 9.20	Důležitost přínosu spojitých proměnných do klastru 1; interval spolehlivosti
Graf 9.21	Důležitost přínosu jedné spojitě proměnné do všech klastrů

## **Přílohy**

## Příloha 1 Vniřní organizační struktura ESS

V rámci European Social Survey pracuje na přípravě výzkumných šetření několik orgánů. Na vrcholu stojí Centrální koordinační tým (*Central Coordinating Team CCT*; Centre for Comparative Social Survey UK), který je odpovědný za design a koordinaci ESS. Centrální multinárodnostní koordinační tým funguje ve složení Roger Jowell (NCSR, UK), Jaak Billiet (University of Leuven, Belgie), Bjørn Henrichsen (NSD, Norsko), Peter Möhler (ZUMA, Německo), Willem Saris (ESADE, Španělsko), Ineke Stoop (SCP, Nizozemí).

Jako odborný poradní orgán je mu k dispozici Vědecký poradní výbor (*Scientific Advisory Board*) pod vedením Maxe Kaase, což je mezinárodní výbor složený z odborníků na kvantitativní výzkumná šetření z účastnicích se zemí (národní koordinátoři), kteří jednotlivě dohlíží na šetření v dané zemi. Dalšími členy Vědeckého poradního orgánu jsou zástupce Evropské komise a zástupce Evropské nadace pro vědu. K dispozici je Centrálnímu koordinačnímu týmu také tým specialistů nazvaný Panel pro výběry (*Sampling panel*) pod odpovědností Sabin Häder (Německo), který s národními koordinátory konzultují a schvalují design výběru v účastnicích se zemích. Komise pro překlady (*Translation Taskforce*) je vedena Janet Harkness (Německo) a představuje další tým specialistů na proces překladu ekvivalentních výzkumných nástrojů ze zdrojové verze dotazníku do národních jazyků. Mezi všemi uvedenými orgány ve struktuře ESS ještě funguje Metodologická skupina (*Methods Group*) vedena Denise Lievesley (UK), která dohlíží na technické metodologické aspekty výzkumu a vyvíjí technické inovace [Jowell, ESS brochure].

Příprava mezinárodního datového souboru, jeho archivace, dokumentace a distribuce jsou svěřeny norskému datovému archivu NSD (Norwegian Social Science Data Services). Ten zveřejňuje první verzi spojeného datového souboru ze všech účastnicích se zemí vždy jeden rok po oficiálním zahájení sběru dat dané vlny šetření. Data jsou dostupná veřejně bez omezení na internetu. Podle pravidel ESS existuje jen jediný společný zdroj pro distribuci dat ESS, a to jak mezinárodních spojených souborů, tak i souborů pro jednotlivé národy.

Ze stránek projektu ESS je možné po registraci získat data a datovou dokumentaci z jednotlivých vln dvěma způsoby, jednak prostřednictvím stránek připravených archivem NSD, a jednak díky systému pro vyhledávání, distribuci a analýzu sociálních dat NESSTAR [Plecitá-Vlachová, Kalvas 2005]. Kompletní přístupná datová dokumentace (*Survey documentation*) zahrnuje: dodatky o populačních statistikách jednotlivých zemí, klasifikacích a kódovacích standardech, seznam proměnných a přehled proměnných a použitých otázek, dokumenty o vážení dat, výběrových plánech, reliabilitě a validitě otázek, hodnocení kvality

dat atp. Součástí dostupné dokumentace je také přehled o sběru dat (*Fieldwork summary*), kde je uveden seznam zemí, doba sběru dat, počet platných rozhovorů a návratnosti. Pokud došlo v nějaké zemi k odchylce či chybě, je u příslušné země ještě uvedena podrobná poznámka. V dokumentaci sběru dat (*Fieldwork documents*) nalezneme hlavní a doplňkový dotazník, karty, instrukce pro tazatele, kontaktní formulář s informacemi o respondentech a pokusech o kontaktování a specifické dokumenty. Všechny dokumenty jsou navíc k dispozici ve všech použitých jazykových verzích.<sup>29</sup>

Výše uvedenou dokumentaci navíc doplňují podrobné údaje o konstrukci otázek do dotazníku a souhrnné kontextové informace o jednotlivých zemích a jejich regionech (*Contextual data and indices*), např. počet obyvatel a HDP, údaje o fertilitě, míře úmrtnosti, počtu uprchlíků atp. (tzv. základní statistiky). Jejich monitoring zajišťuje národní koordinátor v jednotlivých zemích [více <http://www.scp.nl/ess/eventnet/>]. Navíc je do této sekce zahrnut seznam důležitých společenských, politických a ekonomických událostí z jednotlivých zemí (včetně mediálních skandálů a přírodních katastrof), které se vztahují k době realizace příslušného šetření a mohou ovlivnit výsledek výzkumu. ESS integruje tato národní kontextová data a data na evropské úrovni do datového souboru (tzv. *event data bank*) z důvodu zesílení jejich analytické schopnosti.

Za Českou Republiku připravilo všechny vlny šetření ESS výzkumné oddělení Hodnotové orientace ve společnosti Sociologického ústavu Akademie věd ČR (SOÚ AV ČR), národním koordinátorem byla Klára Plecítá-Vlachová. Terénní sběr dat provedla v první vlně agentura STEM, v druhé vlně SC&C, třetí vlny se Česká republika neúčastnila a ve čtvrté vlně zajistila sběr agentura Factum Invenio. První vlnu ESS finančně podpořilo Ministerstvo školství, mládeže a tělovýchovy (MŠMT) v rámci programu INGO na podporu české účasti v programech nevládních organizací. Finanční prostředky byly ale uvolněny až na konci roku 2002 a bylo jich méně než se očekávalo, což se odrazilo na přípravě a průběhu sběru dat. Druhou vlnu ESS finančně podpořila Grantová agentura České Republiky (GA ČR), grant byl udělen v rámci standardní soutěže GA ČR na podporu výzkumu. Čtvrtou vlnu ESS finančně podpořil Sociologický ústav AV ČR. Datové soubory ze všech vln a jejich dokumentaci pro mezinárodní archivaci připravuje Sociologický datový archiv SOÚ AV ČR.

---

<sup>29</sup> Viz. <http://ess.nsd.uib.no>